



King's Research Portal

DOI:

[10.1037/hea0000416](https://doi.org/10.1037/hea0000416)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Webster, R. K., Weinman, J., & James Rubin, G. (2016). A systematic review of factors that contribute to nocebo effects. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, 35(12), 1334-1355. <https://doi.org/10.1037/hea0000416>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

A systematic review of factors that contribute to nocebo effects

Rebecca K. Webster, John Weinman, and G. James Rubin

King's College London

Author note

Rebecca K Webster, King's College London, Department of Psychological Medicine,
Weston Education Centre, United Kingdom

John Weinman, King's College London, Institute of Pharmaceutical Science, Franklin-
Wilkins Building, United Kingdom

G. James Rubin, King's College London, Department of Psychological Medicine, Weston
Education Centre, United Kingdom

The research was funded by the National Institute for Health Research Health Protection
Research Unit (NIHR HPRU) in Emergency Preparedness and Response at King's College
London in partnership with Public Health England (PHE). The views expressed are those of
the authors and not necessarily those of the NHS, the NIHR, the Department of Health or
Public Health England.

Correspondence concerning this article should be addressed to Rebecca K Webster,
Department of Psychological Medicine, Weston Education Centre, Cutcombe Road, London
SE5 9RJ, UK. E-mail: rebecca.webster@kcl.ac.uk

Abstract

Objectives: Medication side effects are common, often leading to reduced quality of life, non-adherence and financial costs for health services. Many side effects are the result of a psychologically-mediated ‘nocebo effect’. This review identifies the risk factors involved in the development of nocebo effects.

Methods: Web of Science, Scopus, Medline, PsychINFO, Journals @ Ovid full text, and Global health were searched using the terms “nocebo” and “placebo effect”. To be included, studies must have exposed people to an inert substance and have assessed one or more baseline or experimental factor(s) on its ability to predict symptom development in response to the inert exposure.

Results: 89 studies were included, 70 used an experimental design and 19 used a prospective design, identifying 14 different categories of risk factor. The strongest predictors of nocebo effects were a higher perceived dose of exposure, explicit suggestions that the exposure triggers arousal or symptoms, observing people experiencing symptoms from the exposure, and higher expectations of symptoms.

Conclusions: In order to reduce nocebo induced symptoms associated with medication or other interventions clinicians could: reduce expectations of symptoms, limit suggestions of symptoms, correct unrealistic dose perceptions, and reduce exposure to people experiencing side effects. There is some evidence that we should do this especially for persons with at-risk personality types, though exactly which personality types these are requires further research. These suggestions have a downside in terms of consent and paternalism, but there is scope to develop innovative ways to reduce nocebo effects without withholding information.

Key words: nocebo effect, predictors, symptoms, inert exposure, review

Adverse drug reactions (ADRs) are common (Davies et al., 2009), and can have serious implications in terms of patient well-being and adherence (Ammassari et al., 2001) as well as significant financial costs for health services (Monguio, Otero, & Rovira, 2003; NICE, 2009). However, ADRs are not always related to the physiological action of the medication (Faasse & Petrie, 2013). Only 10.9% of reported ADRs to commonly prescribed drugs are clearly attributable to the medication (de Frutos Hernansanz et al., 1994). It is thought a placebo effect may play a role in the formation of other apparent side effects (Barsky, Saintfort, Rogers, & Borus, 2002). As well as medication side effects, placebo effects have been implicated in symptoms attributed to technological exposures such as electro-magnetic fields (EMF) from mobile phones and Wi-Fi (Baliatsas et al., 2012; Rubin, Cleare, & Wessely, 2008). A placebo effect is the experience of negative symptoms following exposure to an inert substance, which are triggered or exacerbated by psychological mechanisms such as expectations (Kennedy, 1961). The name 'placebo' was created to distinguish between the desirable ('placebo') and undesirable effects of an inert exposure (Hauser, Hansen, & Enck, 2012), although in practice the distinction between undesirable and desirable is not always clear cut. For example increased alertness may be beneficial in some contexts (e.g. prior to an examination) and detrimental in others (e.g. prior to sleep).

Current literature suggests there are three main mechanisms for a placebo effect; misattribution, expectation, and learning. Misattribution theory suggests that people misattribute pre-existing symptoms to the effects of a new exposure (although some authors believe that misattribution does not technically constitute a placebo effect, see Enck, Bingel, Schedlowski & Rief, 2013; Colloca & Miller, 2011). Symptoms are common in everyday life (Petrie, Faasse, Critchon, & Grey, 2014), and although often harmless and short-lived, when people are subjected to a new exposure, symptoms that were present before or occur coincidentally are available to be mistakenly attributed to it (Petrie et al., 2005; Petrie, Moss-

Morris, Grey, & Shaw, 2004). Therefore factors such as high baseline symptoms or high self-awareness may serve as risk factors for nocebo effects resulting from this mechanism. Negative expectations can also mediate nocebo effects (Hahn, 1997), and may in turn arise through explicit suggestions about the effects of an exposure (Jaen & Dalton, 2014; Myers, Cairns, & Singer, 1987), or predisposing factors such as pessimism (Geers, Helfer, Kosbab, Weiland, & Landry, 2005). These negative expectations can make the individual more likely to attend to new or current sensations, and attribute them to the exposure (Barsky et al., 2002). The response expectancy theory suggests that it is also possible for negative expectations to act more directly, with an expectation of, for example anxiety, being itself provoke anxiety thereby directly causing the negative effect that was expected (Kirsch, 1997a,b). The last mechanism, learning, can elicit nocebo effects through association or social observation. For example, if an inert stimulus has been previously paired with a symptom-inducing stimulus (Barsky et al., 2002), which may occur through conscious or non-conscious mechanisms (Stewart-Williams, 2004), or through observing someone else experience symptoms to the same exposure (Vogtle, Barke, & Kroner-Herwig, 2013).

Given the significant costs nocebo effects can have on patient quality of life and health services it is important to develop interventions to minimise these effects from occurring. Many risk factors have been implicated, but no study has systematically reviewed these to identify those which are the strongest predictors of nocebo effects; something that would assist in the development of such interventions. Instead, previous systematic reviews have focused on the magnitude of nocebo effects for a specific symptom, e.g. Petersen et al. (2014) or in clinical trials of experimental medical treatments (Hauser, Bartram, Bartram-Wunn, & Tolle, 2012). One review (Symon, Williams, Adelasoye, & Cheyne, 2015) has provided a preliminary assessment of some of the risk factors involved in nocebo effects. However this “scoping review” identified only 17 papers – a limited subset of the available

literature. To address this gap our systematic review aimed to identify the risk factors involved in the reporting of any symptom in response to an inert exposure. This will allow the identification of factors which appear to be consistent predictors of nocebo effects and aid in the development of evidenced-based interventions to prevent them from occurring in the future.

Methods

Identification of studies

Searches were carried out on 11th December 2014 using the databases: Web of Science, Scopus, Medline, PsychINFO, Ovid, and Global health. The search terms consisted of “nocebo” or “placebo effect”, and where available, searches were limited to studies with a human sample, with review articles restricted. The reference sections of included studies were also examined as well as papers suggested through personal contacts. No grey literature was searched and no temporal constraints were used. The review followed a previously designed, unpublished protocol.

Selection criteria

Studies were eligible for inclusion if they met the following criteria:

- Studied a human population (healthy volunteers, patients or children were allowed).
- Used an experimental or prospective design.
- Used an inert exposure, i.e. containing no pharmacological or physiological active ingredient.
- Assessed factors on their ability to predict symptom reporting, and these factors could be baseline characteristics or experimentally induced.

- Included an outcome of symptom reporting after participants received an inert exposure. Reported symptoms must not have been due to an active exposure (e.g. studies where an inert exposure was applied after an active exposure such as heat stimulation were excluded, as in this case the symptoms would have resulted from the heat stimulation).
- Measured symptoms via self-report or inferred through objective measures (e.g. scratching behaviour). Such symptoms could be somatic, a measure of arousal or mood. Because of the difficulty in defining when an outcome is aversive or beneficial we took an inclusive approach. For example measures of alertness (where an increase could be aversive in some instances) or contentedness (where decreases might be possible) were both included.
- Publish in any language.

Data extraction

For each study included in the review, details relating to 20 issues were extracted. In summary these related to: sample characteristics, methodological design, type of exposure, experimental conditions and/or baseline risk factors, symptom measurement, statistical analysis and results. Any non-english articles were translated. We differentiated between studies that used an experimental or a prospective design in order to easily identify factors implicated in placebo effects that can be manipulated and those that naturally occur at baseline. For a copy of the data extraction sheet used, see Appendix 1.

Quality assessment

Eligible studies using an experimental design were assessed using the Cochrane Collaboration's Risk of Bias tool (Higgins et al., 2011). For prospective studies, the CASPin (1998) critical appraisal tool was used and adapted to give a 'high,' 'unclear' or 'low' risk of

bias score, which were colour coded red, orange and green respectively. Originally the CASP is scored with yes/no answers but this was re-scored to low risk (yes) and high risk (no) as well as including an unclear risk response for when enough information was not provided, similar to the Cochrane Risk of Bias tool. As these tools had no criteria assessing sample size we looked at this separately.

Review process

RKW conducted the database searches and screened the titles and abstracts of articles to assess their potential relevance. Guidance was obtained from GJR if there was any uncertainty as to including an article for full text review. RKW obtained the full articles for those citations that appeared potentially relevant and checked them against the inclusion criteria. If it was unclear whether an article met the inclusion criteria, consensus was sought from GJR and JW. RKW then independently extracted data for each included study and carried out the quality assessment with guidance from GJR. Due to the expected heterogeneity in the studies we did not plan for any meta-analyses and instead we used a narrative synthesis. There is no general consensus on the best way to carry out a narrative synthesis for systematic reviews (Popay et al., 2006). As such we decided to use a weight of evidence approach. To do this, we identified the strength of evidence for each risk factor based on the number of studies investigating each risk factors and their respective quality.

Results

Search results

The database search retrieved 12582 citations. After removing duplicates 6585 citations remained. After screening titles and abstracts, we reviewed the full text of 88 articles relating to 96 studies. Of these, 13 studies were excluded for not investigating any risk factors for the development of symptoms, nine were excluded for using an active exposure and seven were excluded for not measuring symptoms. Sixty-six articles met the inclusion criteria.

Twenty-one additional articles were identified by reference checks of included articles and through personal contacts; resulting in a total of 87 articles. Two articles reported results on two separate studies each (Walach & Schneider, 2009; Winters et al., 2001) and are referred to as 'Exp 1' or 'Exp 2' where necessary, leaving 87 articles reporting on 89 studies. Of these, 70 were experimental (see table 1) and 19 prospective (see table 2). Figure 1 provides a flow diagram of the study selection according to the Preferred Reporting for Systematic Reviews and Meta-analyses statement (Moher, Liberati, Tetzlaff, & Altman, 2009).

Quality assessment

Experimental studies. The quality of experimental studies was poor (see figure 2), with the main problem being a lack of clear reporting. Thirty-six studies neglected to mention how they carried out randomisation while 22 studies were at high risk of bias for failing to mention if participants were randomised or for not using randomisation at all. Due to the unclear reporting of random sequence generation, the risk for allocation concealment bias followed a similar pattern. For blinding of participants and personnel, studies often failed to state if the experimenters were blind to the manipulation that accompanied the exposure, leaving the risk of bias unclear. Only six studies used adequate blinding procedures, with 12 not using blinding at all. Sixty-five studies used self-report measures, as such blinding of the outcome assessment was judged to be unlikely to influence these results. For 52 studies, drop outs were not addressed, or if they were, they typically failed to explain how this affected the results, leaving the risk of bias unclear. Only one study had lodged a protocol in a publically accessible registry prior to the start of recruitment leaving us unable to assess the risk for selective reporting for the remaining studies. As well as this we looked for justification of sample size to assess if each study was adequately powered. Again this was poorly addressed with only 9 out of the 70 studies mentioning that they carried out an a priori sample size calculation.

Prospective studies. The prospective studies performed well against the quality check (see figure 2). All studies addressed a clearly focused issue with a standardised exposure across all participants. Studies often lacked information about how participants were recruited. However, self-report measures were widely used to minimise bias from experimenters. The identification and control of confounding factors was only deemed an issue for six studies that neglected to control for demographic factors such as gender or age and past symptom reporting. The follow up of participants was judged to be appropriate in 16 studies. Regarding the generalisability of the findings, it was often difficult to know if the results could be applied to the population being studied due to the insufficient information about how participants were recruited. In addition, similarly to the experimental studies, justification for sample size was limited with only one study providing an a priori sample size calculation.

A. Experimentally induced risk factors categories

Seventy experimental studies were included that investigated risk factors which fell into 9 different categories as discussed below (further details in supplementary tables 3-11).

Learning. Twenty three studies manipulated different types of learning on symptom reporting finding some evidence for its role in placebo effects. Four of these investigated prior experience of which two lower quality studies found no significant effects (Bayer, Coverdale, Chiang, & Bangs, 1998; Dinnerstein & Halm, 1970). However, Andre-Obadia, Magnin, and Garcia-Larrea (2011) showed that sham rTMS tended to worsen patients' pain when following an active yet unsuccessful rTMS treatment (however caution is required as no statistical test accompanied this finding), and a high quality study by Stegen et al. (1998) found that participants reported significantly more arousal and respiratory symptoms when completing a breathing trial with room air before a breathing trial with carbon dioxide rather than afterwards. As such there is some evidence that prior experience is involved in the

development of placebo effects. Two studies of mixed quality explored the impact of implicit association supporting its role in the placebo effect, finding that drinking sham caffeine in a coffee solution resulted in significantly more alertness, contentedness, and arousal, than drinking sham caffeine in an orange juice solution (Flaten & Blumenthal, 1999; Mikalsen, Bertelsen, & Flaten, 2001). Three studies of high quality investigated learning through the manipulation of social observation, with two finding a significant effect, broadly supporting its role in the placebo effect. Lorber, Mazzoni, and Kirsch (2007) failed to show any main effects of observing a confederate display symptom behaviours after inhaling a sham environmental toxin which they were also exposed to. However, in a similar study, participants who observed a confederate display symptoms had significantly higher symptom ratings after inhalation than participants who did not (Mazzoni, Foan, Hyland, & Kirsch, 2010). Similarly, patients who watched a video of people scratching compared to those who saw a video of people sitting idle had higher itch and scratching behaviour rating after administration of sham histamine (Papoiu, Wang, Coghill, Chan, & Yosipovitch, 2011), no results were reported for the healthy volunteers in this study.

Of the remaining 14 studies, 13 investigated learning by using classical conditioning to pair inert exposures such as odours with CO₂ inhalation before presenting the inert exposures on their own (De Peuter et al., 2005; Devriese, De Peuter, Van Diest, Van de Woestijne, & Van den Bergh, 2006; Devriese et al., 2000; Devriese et al., 2004; Meulders et al., 2010; Van den Bergh, Kempynck, van de Woestijne, Baeyens, & Eelen, 1995; Van den Bergh, Stegen, & Van de Woestijne, 1997, 1998; Van den Bergh et al., 1999; Van Diest et al., 2006; Winters et al., 2001 exp 1 and 2; Winters et al., 2003). Six studies of mixed quality found significant effects of classical conditioning and although seven found no main effect of conditioning on symptom reporting, six of these were of lower quality. As such there is some evidence for the role of classical conditioning in placebo effects, and that this learning effect

can be generalised to new odours (Devriese et al., 2000; Van den Bergh et al., 1997, 1998). However odour type alone, without classical conditioning is not enough to elicit symptoms as demonstrated in this group of studies and the remaining study in this category (Dalton, 1999).

Perceived dose. Six studies manipulated participant perceptions of the dose of the exposure that they received. Four of these found significant effects with three being of higher quality, broadly supporting a link between higher perceived dose and nocebo effects. Only two studies found no significant effects of dose related to decaffeinated coffee consumption (Flaten, Aasli, & Blumenthal, 2003) or taking a sham sedative pill (Jensen & Karoly, 1991). The remaining four all demonstrated significant main effects: Increasing the setting on a sham shock generator increased pain intensity ratings in two studies (Bayer, Baer, & Early, 1991; Bayer et al., 1998), tension scores increased as a function of perceived dose following decaffeinated coffee consumption in one study (Kirsch & Weixel, 1988), and in a final study being told that a sham EMF exposure would be strong resulted in a higher overall symptom scores compared to being told the exposure would be weak (Szemerszky, Koteles, Lihi, & Bardos, 2010).

Self-awareness. Four studies manipulated self-awareness during exposure. Three higher quality studies found no significant effects with only one lower quality study reporting an effect. As such there is little evidence that self-awareness increases the likelihood of a nocebo effect. Both Geers, Helfer, et al. (2005) and Geers, Helfer, Weiland, and Kosbab (2006) showed no significant main effects of instructing participants to attend to any symptoms or sensations they experienced. Using a distraction task also did not have a significant effect on symptom reporting (Van den Bergh et al., 1998). Gibbons, Carver, Scheier, and Hormuth (1979) however, did find a significant main effect, with participants facing a mirror reporting less perceived arousal than participants not facing a mirror following ingestion of a sham drug.

Type of administration. Two studies of mixed quality tested whether type of administration affects symptom reporting, finding no evidence for a link with placebo effects. There was no difference in symptom reporting between a sham pill and either a saline injection (Goldman, Witton, & Scherer, 1965) or sham acupuncture (Kaptchuk et al., 2006).

Verbal suggestions on performance. Three studies manipulated verbal suggestions about the effect an inert exposure would have on performance. Two higher quality studies found no significant effects with only one lower quality study reporting an effect. As such there is little evidence that suggesting an exposure impairs performance increases the likelihood of a placebo effect. Both Harrell and Juliano (2009) and Nevelsteen, Legros, and Crasson (2007) found no significant main effects of suggesting sham coffee or sham EMF would enhance or impair performance on a task on any of their symptom measures, respectively. However, smokers told that a sham cigarette would impair performance had significantly more craving symptoms than those who were told it would enhance performance (Harrell & Juliano, 2012).

Verbal suggestions of likelihood of exposure. Nine studies manipulated suggestions about the likelihood that an exposure would occur. All studies were of higher quality with four finding significant effects and five finding non-significant effects. In other words, there was mixed evidence for the role of likelihood suggestions in placebo effects. The studies used a mixture of conditions in which participants were either told they would receive an active exposure (deception), might receive an active or inactive exposure (double-blind), would receive an inactive exposure (open) or nothing (control). Five of the studies found no significant main effects (Geers, Helfer, et al., 2005; Geers et al., 2006; Ossege et al., 2005; Walach, Schmidt, Dirhold, & Nosch, 2002; Walach & Schneider, 2009 exp 1). Geers, Wellman, Fowler, Rasinski, and Helfer (2011) however found that participants reported significantly more side effects in response to a sham pill when given deceptive information,

compared to double-blind or control information. In addition, participants given deceptive or double-blind suggestions had a significantly higher increase in alertness following ingestion of sham coffee (Kirsch & Weixel, 1988) and a significantly higher number of adverse events following a sham weight loss supplement (Tippens et al., 2014) than participants in the control condition. For Walach, Schmidt, Bihr, and Wiesch (2001) participants told they would receive an inactive exposure scored higher on general wellbeing than those who received no substance or instruction.

Verbal suggestions of arousal. Sixteen studies manipulated suggestions about the effect an inert exposure would have on arousal. Thirteen studies showed a significant effect, with 10 of these being of higher quality. This strongly supports a link with placebo effects. Only three studies revealed no main effects (Brondeur, 1965; Kuenzel, Blanchette, Zandstra, Thomas, & El-Dereby, 2012; Penick & Fisher, 1965). The remaining 13 all demonstrated significant effects. Participants given stimulant suggestions compared to sedative suggestions had higher tension scores and were more lively after administration of a sham drug (Flaten, Simonsen, & Olsen, 1999; Mrna & Skrivaneck, 1985), and had higher scores of stress, arousal, alertness, friendliness and aggressiveness, and lower fatigue scores after ingestion of an inert drink (Dinnerstein & Halm, 1970; Flaten, 1998; Slanska, Tikal, Hvizdosova, & Benesova, 1974). Higuchi, Shoji, and Hatayama (2002) demonstrated lower stress and stimulant symptoms for participants given relaxing suggestions compared to no information for lavender and jasmine fragrances respectively. Goldman et al. (1965) found that more patients reported suggested drug effects in a sedative condition than in a stimulant condition. The remaining studies found a significant increase in caffeine related symptoms (Geers, Weiland, Kosbab, Landry, & Helfer, 2005; Lotshaw, Bradley, & Brooks, 1996), and alertness (Schneider et al., 2006; Walach & Schneider, 2009 exp 2) and a significant decrease in calmness (Mikalsen et al., 2001) for participants told they would receive caffeine compared

to participants who were told they would not receive caffeine or who received no beverage. Finally Angelucci and Pena (1997) found that participants given coffee with low arousal expectations had significantly lower alertness compared to participants given coffee with no expectations, high arousal expectations or no coffee at all.

Verbal suggestions of symptoms. Twenty one studies manipulated suggestions about what symptoms to expect from an inert exposure. Thirteen found a significant effect, with 11 of these being of higher quality, broadly supporting a link with placebo effects. Out of the 21 studies, eight reported no significant main effects (Devriese et al., 2006; Devriese et al., 2004; Heatherton, Polivy, & Herman, 1989; Jaen & Dalton, 2014; Schweiger & Parnucci, 1981; Walach et al., 2002; Winters et al., 2003; Withöft & Rubin, 2013). For the remaining 13 studies, Benedetti, Amanzio, Casadio, Oliaro, and Maggi (1997), Wise et al. (2009), Crichton, Dodd, Schmid, Gamble, and Petrie (2014) and Pennebaker and Skelton (1981) found significantly higher symptoms scores for those warned about side effects compared to those not warned after administration of sham treatment, infrasound and ultrasonic noise respectively. Dalton (1999), Neukirch and Colagiuri (2014) and Put et al. (2004) found that participants' symptoms were significantly consistent with the warning they received about an odour, sham sleep medication and sham inhaler respectively. Three studies demonstrated that participants experienced significantly more symptoms when informed about side effects to a sham drug (Gibbons et al., 1979; Zimmermann-Viehoff et al., 2013) or saline eye drops (Gavrylyuk, Ehrt, & Meissner, 2010) compared to being informed it was a placebo. Similarly both Bayer et al. (1991) and Read and Bohr (2014) established significantly higher symptoms scores for those informed they would receive an active compared to an inactive exposure. Colagiuri, McGuinness, Boakes, and Butow (2012) however found the opposite, participants not warned about the side effects experienced more and a greater severity of side effects than those warned about one or four side effects.

Miscellaneous. Six studies looked at factors that did not fit into the above categories. There was no significant effect of manipulating participants to cooperate (Geers, Weiland, et al., 2005) or the experimenters' expectations of participants' symptoms (Walach et al., 2001). However, Faasse, Cundy, Gamble, and Petrie (2013) found that manipulating tablet brand to make participants think they had changed to a generic version resulted in a significantly higher number of symptoms compared to participants told that they were still taking the original branded tablet, although this study was of lower quality than the others in this group. Jensen and Karoly (1991) have shown that manipulating social desirability so that participants think responding to the pill is more socially desirable results in significantly higher symptom scores. Type of breathing has also been shown to affect symptom reporting with normocapnic overbreathing resulting in higher respiratory symptoms compared to spontaneous breathing (Van Diest et al., 2006). Lastly, a conditioned odour results in more symptoms if the odour is presented immediately rather than a week after conditioning trials (Devriese et al., 2000).

B. Baseline risk factors categories

Nineteen prospective studies and also 33 experimental studies which assessed baseline risk factors were included which fell into six different categories as discussed below (further details in supplementary tables 12-17).

Demographics. Twenty-one studies looked at the risk of demographic characteristics, finding no demonstrable evidence for their role in placebo effects. Five of these investigated age and found it did not predict any symptom outcomes (de la Cruz, Hui, Parsons, & Bruera, 2010; Geers, Helfer, et al., 2005; Goetz et al., 2008; Lombardi, Gargioni, Canonica, & Passalacqua, 2008; Witthöft & Rubin, 2013). As four of these studies were of higher quality, this is good evidence that age is not linked with the development of placebo effects. Eighteen studies (Angelucci & Pena, 1997; Casper, Tollefson, & Nilsson, 2001; Geers, Helfer, et al.,

2005; Geers et al., 2011; Goetz et al., 2008; Harrell & Juliano, 2012; Jensen & Karoly, 1991; Liccardi et al., 2004; Lombardi et al., 2008; Lorber et al., 2007; Mazzoni et al., 2010; Papoiu et al., 2011; Read & Bohr, 2014; Strohle, 2000; Van den Bergh et al., 1997, 1998; Witthöft & Rubin, 2013) looked at gender and only four reported significant results suggesting women are more susceptible to placebo effects than men (Casper et al., 2001; Liccardi et al., 2004; Strohle, 2000; Szemerszky et al., 2010). Of the remaining 14 showing non-significant effects, 12 were of high quality suggesting there is very little evidence for the role of gender in placebo effects. The effects of level of education (de la Cruz et al., 2010; Witthöft & Rubin, 2013) were equivocal in two high quality studies, whereas employment (Drici, Raybaud, Delunardo, Iacono, & Gustovic, 1995), was not a significant predictor.

Clinical Characteristics. Fourteen studies investigated clinical characteristics, finding mixed evidence for a link with placebo effects. Six studies of high quality looked at the effect of baseline symptom scores, finding mixed evidence for a link with placebo effects. Two found no significant effects (Andre-Obadia et al., 2011; Casper et al., 2001). For the other four, results were mixed. Danker-Hopfe, Dorn, Bornkessel, and Sauter (2010) and de la Cruz et al. (2010) found that higher symptom scores at baseline predicted higher symptom scores after exposure to sham EMF and treatment respectively, whereas Flaten et al. (2003) and Goetz et al. (2008) found the opposite after drinking decaffeinated coffee and taking sham medication for Parkinson's respectively. Six studies of high quality looked at the effect of type of clinical condition, with five finding a significant effect. They showed that suffering from a condition that is exacerbated by the suggested sham exposure significantly increased symptom reporting compared to healthy volunteers, strongly supporting a link with placebo effects. Nevelsteen et al. (2007) found that depression did not predict symptoms in response to a sham magnetic field. However, Papoiu et al. (2011), Strohle (2000), De Peuter et al. (2005) and Bogaerts et al. (2010) showed that suffering from atopic dermatitis, panic

disorder, asthma or medically unexplained dyspnea resulted in significantly more symptoms in response to sham histamine, sham panic disorder trigger, sham inhaler and breathing trials with room air, respectively, compared to healthy volunteers. In addition Szemerszky et al. (2010) found that the level of perceived sensitivity to EMFs was positively correlated with symptom scores after sham EMF exposure. The remaining two studies looked at previous drug reactions finding weak evidence for a link with placebo effects. Lombardi et al. (2008) found no significant effects of type or severity of previous drug reaction on symptoms in response to a sham allergen pill. However, a higher quality study by Mrna and Skrivaneck (1985) found the reaction to another sham drug was significantly correlated with perceived drug effect.

Expectations. Thirteen studies looked at the effect of participant expectations on symptom reporting, broadly supporting a link with placebo effects. Eleven of these studies looked at participants' symptom expectations, of which five higher quality studies revealed no significant effects (Angelucci & Pena, 1997; Molcan & et al., 1982; Walach et al., 2001; Walach & Schneider, 2009 exp 1 and 2). The remaining six studies demonstrated that expectations of symptoms significantly predicted (Fillmore & Vogel-Sprott, 1992; Köteles & Babulka, 2014; Vase et al., 2013) or correlated (De Peuter et al., 2005; Flaten et al., 2003; Szemerszky et al., 2010) with symptom reporting. Five of these studies were of higher quality therefore broadly supporting a link with placebo effects. Three studies also looked at expectations in terms of the substance taken finding weak evidence for its role in placebo effects. Link, Haggard, Kelly, and Forrer (2006) found that participants who believed they had taken an active pill reported more symptoms than those who thought they had taken a sham pill, however this was a low quality study. Higher quality studies by Bayer et al. (1998) and Walach et al. (2001) also investigated this but found no significant effects.

Anxiety. Nine studies looked at the influence of anxiety on symptom reporting, finding weak evidence for a link with placebo effects. Six studies of mixed quality looked at state anxiety (Bogaerts et al., 2010; Link et al., 2006; Molcan & et al., 1982; Nevelsteen et al., 2007; Szemerszky et al., 2010; Witthöft & Rubin, 2013) but only Nevelsteen et al. (2007) found a significant effect, with state anxiety predicting physical symptom scores. Molcan and et al. (1982) and Nevelsteen et al. (2007) found no significant effects of trait anxiety. Angelucci and Pena (1997) found combined state and trait anxiety scores significantly predicted anxiety, but did not report results for state and trait anxiety separately. However no such effect of combined state and trait anxiety was found on symptom reporting to an odour (Van den Bergh et al., 1997), although this was a lower quality study. Finally, a high quality study by Danker-Hopfe et al. (2010) found that anxiety towards a local base station predicted subjective sleep quality after sham EMF exposure.

Personality. Twenty-two studies looked at different aspects of personality as predictors of symptoms. Twelve studies showed significant effects of personality of which only three were of low quality as such finding evidence broadly supporting a link with placebo effects. There were no significant effects of suggestibility (Angelucci & Pena, 1997), sensitivity to anxiety (Nevelsteen et al., 2007), restraint (Heatherton et al., 1989), or social desirability (Link et al., 2006; Put et al., 2004; Stegen, Van Diest, Van de Woestijne, & Van den Bergh, 2000). However, studies did show significant effects of the following on at least one symptom outcome: Type A personalities reported more side effects than type B (Drici et al., 1995); pain catastrophizing positively correlated with side effect reports (Sullivan, Lynch, Clark, Mankovsky, & Sawynok, 2008); blunting behaviour predicted symptom reporting (Van den Bergh et al., 1997); positive affect and vigilance predicted symptom scores (Nevelsteen et al., 2007); “frail and submissive” personality correlated with the exposures perceived effect (Slanska et al., 1974); somatisation and motivation predicted symptom score

(Szemerszky et al., 2010); and modern health worries and somatosensory amplification predicted symptom scores (Witthöft & Rubin, 2013). There was mixed evidence for the role of negative affect (Bogaerts et al., 2010; De Peuter et al., 2007; De Peuter et al., 2005; Devriese et al., 2000; Devriese et al., 2004; Nevelsteen et al., 2007; Put et al., 2004; Stegen et al., 1998; Stegen et al., 2000; Van den Bergh et al., 1995), neuroticism (Mazzoni et al., 2010; Davis, Ralevski, Kennedy, & Neitzert, 1995), and pessimism (Geers, Helfer, et al., 2005; Szemerszky et al., 2010).

Miscellaneous. Thirteen studies looked at baseline factors which did not fit into the above categories. These included caffeine consumption (Geers, Weiland, et al., 2005; Geers et al., 2011), olfactory sensitivity (Dalton, 1999), perceived cue odour (Devriese et al., 2004), visibility of a mobile phone base station and pre-occupation with EMF (Danker-Hopfe et al., 2010), geographical site of enrolment (Goetz et al., 2008), hospital centre (Liccardi et al., 2004), stress experienced whilst wearing a helmet delivering sham EMF (Nevelsteen et al., 2007), ability to predict which odour produced the most symptoms (Meulders et al., 2010) and risk perception (Nevelsteen et al., 2007), which had no significant effects. Köteles and Babulka (2014) however found that odour pleasantness predicted perceived change in alertness for eucalyptus oil. In addition, odour reactivity predicted symptom responding to odours (Dalton, 1999) and high regard for medications positively correlated with perceived drug effect (Goldman et al., 1965). Mazzoni et al. (2010) found that if the gender of the model matched the participant this predicted symptom development in social observation studies. Nevelsteen et al. (2007) found that less comfort under the helmet delivering the sham EMF predicted symptoms. Finally Wendt et al. (2014) reported that significantly more symptoms were reported in val/val homozygous carriers compared to val 158/Met 18 and Met/Met 158 homozygous carriers after sham treatment.

C. Interactions between risk factor categories

As well as investigating the main effects of each risk factor, some studies assessed the interactions between risk factors, as displayed in the last column of tables 3-19. Those risk factors which were implicated often in these interactions were factors such as ‘likelihood suggestion’ which interacted with: ‘pessimism’ - participants given deceptive suggestions report more symptoms compared to those told it was an inactive pill, if they were pessimists (Geers, Helfer, et al., 2005); ‘self-awareness’ - participants given deceptive suggestions reported more symptoms when asked to monitor their bodily sensations (Geers et al., 2006); and ‘perceived dose’ - tension increased with increasing coffee dose for those given deceptive suggestions, but decreased with increasing coffee dose when given double-blind suggestions (Kirsch & Weixel, 1988).

In addition, ‘classical conditioning’ showed interactions with ‘odour’; pairing an odour with CO₂ elicited symptoms to the odour alone, only if the odour was foul smelling (Devriese et al., 2000; Van den Bergh et al., 1995; Van den Bergh et al., 1997; Winters et al., 2003). This interaction between ‘classical conditioning’ and ‘odour’ was also found to more likely occur among people with high ‘negative affect’ (Devriese et al., 2000) and those manipulated to have higher ‘self-awareness’ (Van den Bergh et al., 1998). Negative affect also interacted with ‘symptom suggestions’, with higher obstruction and dyspnea symptom scores after suggestions of bronchoconstriction compared to bronchodilation for a sham inhaler if participants had high negative affect (Put et al., 2004). An interaction was also found with ‘prior experience’, with high negative affect participants reporting more arousal and symptoms on the whole to a room-air breathing trial when this preceded rather than followed a CO₂ breathing trial (Stegen et al., 1998).

As well as interacting with negative affect, symptom suggestions interacted with other factors. These included: ‘self-awareness’, participants reported more symptoms when told they were taking an active drug with side effects if they were not facing a mirror (Gibbons et

al., 1979); ‘odours’, more symptom reports following suggestion of symptoms if the odour was unpleasant (Dalton, 1999); ‘classical conditioning’, higher total, respiratory, cardiac and unclassified symptom scores following exposure to an odour previously paired with CO₂ if participants received symptom suggestions (Winters et al., 2003); and ‘state anxiety’, higher total and head/concentration symptoms following symptom suggestions if participants had high anxiety (Witthöft & Rubin, 2013).

Discussion

Summary of main results

From the 89 studies that met our inclusion criteria, 14 categories of risk factor for a nocebo effect were identified, including nine experimentally induced risk factor categories and six baseline risk factor categories. Of these categories, ‘learning/social observation’, ‘perceived dose,’ ‘verbal suggestions of arousal and symptoms’, and ‘baseline symptom expectations’ appeared to be the strongest predictors of nocebo effects. There was some evidence for the role of ‘personality’ in nocebo effects; however which facets of personality are more strongly linked with nocebo effects needs further research. In addition, although not strong predictors on their own, learning/classical conditioning, likelihood suggestion, self-awareness and negative affect consistently interacted with other risk factors.

Given the proposed psychological mechanisms behind nocebo effects it is perhaps unsurprising that these factors have been consistently identified in the literature. Specifically looking at the expectation mechanism, it is intuitive that verbal suggestions of symptoms can generate expectations of these effects leading to symptom reporting. In support of this, participants’ own baseline expectations can trigger symptoms, while perceived dose presumably affects symptom reports through a mediating effect of expectations, with a higher dose associated in a participant’s mind with a stronger effect. This could also explain the

significance of medication brand, with branded medication being generally expected by the public to be better quality than generic unbranded medication and therefore less likely to cause side effects (Faasse et al., 2013). Expectations could also explain why four studies which measured symptom reports both for pre-warned and non-warned symptoms found stronger effects for symptoms that had previously been suggested (Faasse et al., 2013; Gibbons et al., 1979; Lorber et al., 2007; Mazzoni et al., 2010). It also explains why no effect was found for performance suggestions, as this should not directly influence expectations of symptoms from the exposure.

It is important not to over-emphasise the nature of our results with respect to expectation, however. In particular, it was striking that type of administration and verbal suggestions of the likelihood of exposure did not appear to be relevant despite both supposedly raising expectations of symptoms. Possibly, the influence of these factors on expectations is weaker than might be thought. Alternatively, methodological factors may account for the lack of effect. For example, both studies assessing type of administration used patient samples (Goldman et al., 1965; Kaptchuk et al., 2006). Given their greater experience with medical procedures, merely changing an intervention from a pill to an injection may not have triggered a substantial change in expectations. For three of the likelihood suggestion studies (Walach et al., 2001; Walach et al., 2002; Walach & Schneider, 2009 exp 1) it was suggested that the absence of an effect could have been because of cultural differences, with the caffeine effect stereotype not as strong in Germany as it is in the USA.

The overall support for the role of expectations identified in our review still allows for at least two ‘sub-mechanisms’ to exist. The first is a role for attentional bias and symptom detection (Hahn, 1997). The second is a more direct effect, where-by expectations affect emotional state (Kirsch, 1997b; Stewart-Williams, 2004). For example Kirsch (1997b) pointed out that the expectation of anxiety is likely to be anxiety provoking, thereby directly

causing the outcome. This could explain the strong results seen for manipulating verbal suggestions of arousal on symptom reporting, as the expectation of arousal or relaxation is itself likely to be arousing or relaxing. However, there does need to be a degree of caution in interpreting these results on arousal as they could be interpreted as part of the placebo response.

With regards to misattribution as a mechanism, the evidence from the studies that investigated self-awareness as a risk factor did not support this, with the two most directly relevant studies that instructed participants to monitor for any sensations failing to find an effect. Equally, for the six studies investigating the effect of baseline symptoms on symptom reporting the results were mixed providing inconclusive support for misattribution. However five studies (Bogaerts et al., 2010; De Peuter et al., 2005; Papoiu et al., 2011; Strohle, 2000; Szemerszky et al., 2010), showed that suffering from a condition with symptoms similar to those being induced was a predictor of symptom reporting. As such while the mechanism remains plausible, further evidence is required to clarify its importance.

For the learning mechanism support was found from studies investigating the risk factor ‘association’, with the taste of decaffeinated coffee being enough to elicit caffeine related symptoms (Flaten & Blumenthal, 1999; Mikalsen et al., 2001). For prior experience, the results were weak but this could have been due to a lack of experience as this manipulation was typically a one off event. However, there was evidence for the role of social observation, with two out of three studies showing a significant effect. In addition, support for learning was seen in the studies using classical conditioning, which involved a number of trials. Almost half of the studies showed that conditioning CO₂ inhalation with any odour is enough to elicit symptoms to the odour itself, and a reliable finding amongst the studies was that this was especially the case if the odour was unpleasant.

For baseline risk factors, we found no evidence of any effects of gender. However, since conducting the literature search, one additional study that would have met the inclusion criteria has become apparent and which is relevant here. This study by Faasse, Grey, Jordan, Garland, and Petrie (2015) investigated the risk factor of observing a female confederate display symptoms, demonstrating a significant effect on symptom reporting in females. It is interesting to note that Lorber et al. (2007), who also studied social observation, also only found a significant effect in females. One possibility is that it may be something inherent to social observation that makes females more vulnerable to nocebo effects. Other demographic factors such as age, employment status or level of education were also not risk factors. Interestingly, anxiety did not come out as a strong predictor despite the role it could play through misattribution (generating physical symptoms that are available to be misattributed) and expectations (apprehension of symptoms). One possible explanation for this advanced by Szemerszky et al. (2010) is that scores of anxiety could reach a ceiling effect due to advance information about the risks of taking part in the study. For other baseline risk factors, many different types of personality were implicated such as: type A personality (Drici et al., 1995), lower positive affect, vigilance (Nevelsteen et al., 2007), pessimism, motivation to cooperate, somatisation, somatosensory amplification, modern health worries (Szemerszky et al., 2010; Witthöft & Rubin, 2013), and neuroticism (Davis et al., 1995). A lack of consistency in the personality traits studied makes it difficult to interpret these findings, but many would seem to fit with expectation and / or misattribution mechanisms.

Nocebo effects have occasionally been referred to as the ‘evil twin’ of placebo effects. If true, one would expect the risk factors for a nocebo effect to be the inverse of the risk factors for a placebo effect. At a first look the mechanisms supported in our review do appear to be similar to those previously identified for placebo effects, albeit acting in the opposite direction. For example, the expectancy mechanism has been implicated for placebos through

factors such as verbal suggestions, and participants' own baseline expectations which lead to positive expectations for pain or symptom relief (Benedetti et al., 2003; Kam-Hansen et al., 2014; Price et al., 1999; Vits et al., 2013). In addition, learning mechanisms such as prior experience of pain relief, social observation or conditioning people to experience pain relief results in subsequent placebo responses (Colloca & Benedetti, 2006; Colloca & Benedetti, 2009; Suchman & Ader, 1992). It also seems that opposite personality characteristics also predict placebo responding e.g. optimism (Geers, Kosbab, Helfer, Weiland, & Wellman, 2007) as opposed to pessimism. One notable exception, however, would be the misattribution of pre-existing symptoms, as logically this can only be relevant for nocebo: one cannot misattribute the absence of pre-existing symptoms to an exposure. However it is possible one could misattribute and fixate on a coincidental decline in symptoms after taking a sham tablet, and misattribute their improved wellbeing to the tablet.

Quality of original research

It is possible that some of our conclusions may be due to differences in quality between those studies that found an effect and those that did not. We did not observe any clear trend for lower quality studies to report more or fewer significant results than higher quality studies. However, on the whole the quality of the studies included in this review was limited due to poor reporting of key issues in experimental research such as randomisation, allocation concealment, blinding, and not registering a study protocol prior to initiating recruitment. Prospective studies had fewer quality concerns, however given that experimental studies allow the control of more variables the results of these have more weighting than those from the prospective studies. It is also worth noting that almost half of studies did not mention receiving ethical approval. In an area of research requiring deception, or at least, withholding information in order to deliberately cause symptoms, this is surprising. There is scope for future researchers to improve the methodological rigour of this field. Another

surprising limitation of many of the studies included in this review was the lack of a priori sample size calculations. Only 10 out of 89 studies included in this review mentioned carrying out a sample size calculation in order to make sure the sample was adequately powered to test their research question(s). As such we could not assess the quality of studies based on their sample size in the large majority cases. Although it would have been useful to score each study for their strength of evidence, due to this lack of clear reporting and the heterogeneity across studies it was too hard to quantify the strength of each study using the same scale.

Quality of this review

A strength of this review is that we did not include studies in which participants were exposed to an active exposure capable of eliciting symptoms through physiological mechanisms (e.g. experiments altering the information given to participants about a genuine medication). Such studies do not assess the pure nocebo effect, described as the undesirable effects experienced from an inert exposure (Kennedy, 1961) and can prove more difficult to interpret (Neukirch & Colagiuri, 2014).

Our search resulted in a large number of results. As the term ‘nocebo’ is still not widely used and may be preferentially used by those studies identifying a significant increase in symptoms in their participants, we deliberately adopted a broader search strategy than that used in previous reviews, e.g. Petersen et al. (2014). Despite this, it is not certain every study that met the inclusion criteria has been included, especially as nearly a quarter of included studies were identified through personal contacts. This inconsistent use of terminology makes the nocebo literature difficult to search and will continue to limit reviews in this area. We could have included terms such as ‘adverse effects or negative outcome’ in the search strategy but the number of results would be unmanageable as it would include many clinical trials that would not meet our inclusion criteria. On medline alone, such search terms return

over 97,000 results. This is also one of the reasons why we did not simply use ‘placebo’ as one of the search terms – every study which described itself as “placebo-controlled” would be returned.

In addition to limitations resulting from our search strategy, it is possible that some studies could have been falsely rejected after title and abstract screening (e.g. the main purpose of the study may have been on the placebo effect and therefore only placebo and not nocebo findings were reported in the abstract). We suspect that this is unlikely to have occurred often, however. In order to have been included such studies would have had to a) manipulated factor(s) in order to affect nocebo responding or b) looked at baseline measures as predictors of nocebo responding, which many do not do. Many studies which looked at the placebo effect passed through abstract screening as they mentioned participants experiencing negative symptoms or patients feeling worse after placebo exposure. However, going through the full manuscript the majority of these studies would not explore the possible reasons why, e.g. baseline predictors. Therefore we feel this is not something to be too concerned about.

In addition studies published in non-European languages may have been less likely to have been identified as well as studies that were not reported in the conventional peer-reviewed literature.

Other limitations of the review reflect the way we grouped the results. We aggregated studies based on the independent variable. Because of this and due to the fact that there are no direct replications each risk factor grouping contains several different outcomes. It is possible that an interaction exists between independent and dependent variables: for example, some outcomes may be more susceptible to the effects of changes in expectations than others. Unfortunately, we did not have enough data to explore this in depth.

Similarly as this review focused on identifying all the possible risk factors of nocebo effects that have been investigated in the literature, we included studies with different research populations, e.g. students, healthy volunteers and patients. As such there could be differences between the groups in terms of which mechanisms are more likely to be at play. For example, it is likely the misattribution mechanism is more important for the development of nocebo effects in patient samples than healthy volunteers. However, looking at studies that had a patient sample we should interpret the results of those that just focused on baseline disease measures as support of the misattribution mechanism with caution. These studies did not measure actual baseline symptoms or emotions which are more likely to be subject to the misattribution mechanism, rather than disease status.

Finally, the interaction between the mechanisms, outcomes and mode of delivery may also be important, but could not be explored in detail given the data available to us. For example, different forms of sham intervention e.g. sham tablets vs sham caffeine vs sham EMF, may be more or less likely to trigger certain psychological mechanisms, and be more or less likely to affect certain outcomes, e.g. Szermersky, Dömötör, Berkes and Koteles (2016).

Implications for clinical practice and research

Our results suggest clinicians keen to reduce side effects induced by any nocebo effect associated with their interventions could: 1) identify patient expectations of the adverse effects of an intervention and provide reassurance if these seem excessive; 2) avoid giving suggestions of side effects associated with the intervention; 3) down-play the dose that is being provided; 4) reduce patient exposure to other patients experiencing side effects. Wells and Kaptchuk (2012) suggest the use of contextualised informed consent, whereby doctors should identify high risk patients and tailor the medication side effect information so that these patients only receive drug specific side effect information, which is less susceptible to the nocebo response. Our review supports this and suggests that such tailoring may be

especially required for those who have at-risk personality types. Clearly, these suggestions also have a downside, however, as they reduce informed consent and patient autonomy by restricting the information that is being provided. Alternative ways to reduce nocebo effects while maintaining the ability of a patient to give full informed consent are required. There is scope for researchers to develop innovative ways to reduce nocebo effects that does not require withholding of information. This has been shown by Crichton and Petrie (2015) who found that informing participants about nocebo effects effectively reduced symptoms to infrasound noise. In addition Bingel (2014) provides some suggestions on how to avoid nocebo effects which are supported by this review such as improving the communication in patient information leaflets to make them more patient-orientated and reduce negative expectations of potential adverse effects.

Additional research should also aim to replicate risk factors which have so far received limited research, such as the more rarely investigated personality characteristics. It would also be advisable to look again at the risk factor ‘type of administration’ in a healthy volunteer sample and to assess this manipulation on expectations to explore possible mechanisms. It is also time for authors to use consistent terminology allowing easier identification of papers, and to enhance the quality of their research in this area. Simple acts such as being more explicit about randomisation and blinding procedures and publishing protocols will enhance the transparency of the research in this area whilst also helping to alleviate some of the controversy surrounding nocebo research.

Conclusions

This review found that there is a mix of factors which predict whether someone will experience a nocebo effect. Given the implications nocebo effects have on patients’ quality of life and the health costs they create, it is important for research to start developing interventions to prevent nocebo effects from occurring whilst still trying to uphold informed

consent. This systematic review provides a useful starting point for researchers to develop evidenced based interventions designed to negate nocebo effects, whilst also highlighting areas that need further investigation and improvement.

Acknowledgements

We would like to thank the reviewers for making such detailed and insightful comments during the peer review process.

References

- Ammassari, A., Murri, R., Pezzotti, P., Trotta, M. P., Ravasio, L., De Longis, P., . . .
- Antinori, A. (2001). Self-Reported Symptoms and Medication Side Effects Influence Adherence to Highly Active Antiretroviral Therapy in Persons With HIV Infection. *Journal of Acquired Immune Deficiency Syndromes*, 28, 445-449.
- Andre-Obadia, N., Magnin, M., & Garcia-Larrea, L. (2011). On the importance of placebo timing in rTMS studies for pain relief. *Pain*, 152(6), 1233-1237. doi: 10.1016/j.pain.2010.12.027
- Angelucci, L., & Pena, G. (1997). Expectations as explanatory variable of the placebo effect. *Revista Interamericana De Psicologia*, 31(1), 109-132.
- Baliatsas, C., Van Kamp, I., Bolte, J., Schipper, M., Yzermans, J., & Lebre, E. (2012). Non-specific physical symptoms and electromagnetic field exposure in the general population: Can we get more specific? A systematic review. *Environmental International*, 41, 15-28.
- Barsky, A. J., Saintfort, R., Rogers, M. P., & Borus, J. F. (2002). Nonspecific medication side effects and the nocebo phenomenon. *JAMA*, 287(5), 622-627.
- Bayer, T. L., Baer, P. E., & Early, C. (1991). Situational and psychophysiological factors in psychologically induced pain. *Pain*, 44(1), 45-50.

- Bayer, T. L., Coverdale, J. H., Chiang, E., & Bangs, M. (1998). The role of prior pain experience and expectancy in psychologically and physically induced pain. *Pain*, 74(2-3), 327-331.
- Benedetti, F., Amanzio, M., Casadio, C., Oliaro, A., & Maggi, G. (1997). Blockade of nocebo hyperalgesia by the cholecystokinin antagonist proglumide. *Pain*, 71(2), 135-140. doi: 10.1016/s0304-3959(97)03346-0
- Benedetti, F., Maggi, G., Lopiano, L., Lanotte, M., Rainero, I., Vighetti, S., & Pollo, A. (2003). Open versus hidden medical treatments: the patient's knowledge about a therapy affects the therapy outcome. *Prevention and Treatment*, 6(1). doi: 10.1037/1522-3736.6.1.61a
- Bingel, U. (2014). Avoiding Nocebo Effects to Optimize Treatment Outcome. *Journal of the American Medical Association*, 312(7), 693-694. doi: 10.1001/jama.2014.8342
- Bogaerts, K., Van Eylen, L., Wan, L., Bresseleers, J., Van Diest, I., De Peuter, S., . . . Van den Bergh, O. (2010). Distorted symptom perception in patients with medically unexplained symptoms. *Journal of Abnormal Psychology*, 119(1), 226-234.
- Brodeur, D. (1965). The effects of stimulant and tranquilizer placebos on healthy subjects in a real life situation. *PSYCHOPHARMACOLOGIA*, 7(6), 444-452.
- Casper, R. C., Tollefson, G. D., & Nilsson, M. E. (2001). No gender differences in placebo responses of patients with major depressive disorder. *Biological Psychiatry*, 49(2), 158-160. doi: 10.1016/s0006-3223(00)00966-5
- Critical Appraisal Tools. (1998). <http://www.caspinternational.org/?o=1012>.
- Colagiuri, B., McGuinness, K., Boakes, R. A., & Butow, P. N. (2012). Warning about side effects can increase their occurrence: an experimental model using placebo treatment for sleep difficulty. *Journal of Psychopharmacology*, 26(12), 1540-1547. doi: 10.1177/0269881112458730

- Colloca, L., & Benedetti, F. (2006). How prior experience shapes placebo analgesia. *Pain*, 124, 126–133.
- Colloca, L., & Benedetti, F. (2009). Placebo analgesia induced by social observational learning. *Pain*, 144(1-2), 28-34. doi: 10.1016/j.pain.2009.01.033
- Colloca, L., & Miller, F. G. (2011). The Nocebo Effect and Its Relevance for Clinical Practice. *Psychosomatic Medicine*, 73(7), 598-603. doi: 10.1097/PSY.0b013e3182294a50
- Crichton, F., Dodd, G., Schmid, G., Gamble, G., & Petrie, K. J. (2014). Can expectations produce symptoms from infrasound associated with wind turbines? *Health Psychology*, 33(4), 360-364.
- Crichton, F., & Petrie, K. J. (2015). Health complaints and wind turbines: The efficacy of explaining the nocebo response to reduce symptom reporting. *Environmental Research*, 140, 449-455. doi: <http://dx.doi.org/10.1016/j.envres.2015.04.016>
- Dalton, P. (1999). Cognitive influences on health symptoms from acute chemical exposure. *Health Psychology*, 18(6), 579-590.
- Danker-Hopfe, H., Dorn, H., Bornkessel, C., & Sauter, C. (2010). Do mobile phone base stations affect sleep of residents? Results from an experimental double-blind sham-controlled field study. *American Journal of Human Biology*, 22(5), 613-618.
- Davies, E. C., Green, C. F., Taylor, S., Williamson, P. R., Mottram, D. R., & Pirmohamed, M. (2009). Adverse Drug Reactions in Hospital In-Patients: A Prospective Analysis of 3695 Patient-Episodes. *Plos One*, 42(2). doi: 10.1371/journal.pone.0004439
- Davis, C., Ralevski, E., Kennedy, S. H., & Neitzert, C. (1995). The Role of Personality Factors in the Reporting of Side Effect Complaints to Moclobemide and Placebo: A Study of Healthy Male and Female Volunteers. *Journal of Clinical Psychopharmacology*, 15(5), 347-352.

- de Frutos Hernansanz, M. J., Lázaro, D. A., Llinares, G. V., Azpiazu, G. M., Serrano, V. A., & López de Castro, F. (1994). Adverse reactions to drugs in a health center. *Atención Primaria*, 14(5), 783-786.
- de la Cruz, M., Hui, D., Parsons, H. A., & Bruera, E. (2010). Placebo and Nocebo Effects in Randomized Double-Blind Clinical Trials of Agents for the Therapy for Fatigue in Patients With Advanced Cancer. *Cancer*, 116(3), 766-774. doi: 10.1002/cncr.24751
- De Peuter, S., Put, C., Lemaigre, V., Demedts, M., Verleden, G., & Van den Bergh, O. (2007). Context-evoked overperception in asthma. *Psychology & Health*, 22(6), 737-748.
- De Peuter, S., Van Diest, I., Lemaigre, V., Wan, L., Verleden, G., Demedts, M., & Van den Bergh, O. (2005). Can subjective asthma symptoms be learned? *Psychosomatic Medicine*, 67(3), 454-461.
- Devriese, S., De Peuter, S., Van Diest, I., Van de Woestijne, K. P., & Van den Bergh, O. (2006). US-inflation in a differential odor-conditioning paradigm is not robust: Relevance for medically unexplained symptoms. *Journal of Behavior Therapy and Experimental Psychiatry*, 37(4), 314-332. doi: <http://dx.doi.org/10.1016/j.jbtep.2006.03.003>
- Devriese, S., Winters, W., Stegen, K., Van Diest, I., Veulemans, H., Nemery, B., . . . Van den Bergh, O. (2000). Generalization of acquired somatic symptoms in response to odors: a pavlovian perspective on multiple chemical sensitivity. *Psychosomatic Medicine*, 62(6), 751-759.
- Devriese, S., Winters, W., Van Diest, I., De Peuter, S., Vos, G., Van de Woestijne, K., & Van den Bergh, O. (2004). Perceived relation between odors and a negative event determines learning of symptoms in response to chemicals. *International Archives of*

Occupational and Environmental Health, 77(3), 200-204. doi: 10.1007/s00420-003-0488-8

Dinnerstein, A. J., & Halm, J. (1970). Modification of placebo effects by means of drugs: Effects of aspirin and placebos on self-rated moods. *Journal of Abnormal Psychology*, 75(3), 308-314.

Drici, M. D., Raybaud, F., Delunardo, C., Iacono, P., & Gustovic, P. (1995). INFLUENCE OF THE BEHAVIOR PATTERN ON THE NOCEBO RESPONSE OF HEALTHY-VOLUNTEERS. *British Journal of Clinical Pharmacology*, 39(2), 204-206.

Enck, P., Bingel, U., Schedlowski, M., & Rief, W. (2013). OPINION The placebo response in medicine: minimize, maximize or personalize? *Nature Reviews Drug Discovery*, 12(3), 191-204. doi: 10.1038/nrd3923

Faasse, K., Cundy, T., Gamble, G., & Petrie, K. J. (2013). The Effect of an Apparent Change to a Branded or Generic Medication on Drug Effectiveness and Side Effects. *Psychosomatic Medicine*, 75(1), 90-96. doi: 10.1097/PSY.0b013e3182738826

Faasse, K., Grey, A., Jordan, R., Garland, S., & Petrie, K. J. (2015). Seeing is believing: Impact of social modeling on placebo and nocebo responding. *Health Psychology*, 34(8), 880-885.

Faasse, K., & Petrie, K. J. (2013). The nocebo effect: patient expectations and medication side effects. *Postgraduate Medical Journal*, 89, 540-546.

Fillmore, M., & Vogel-Sprott, M. (1992). Expected effect of caffeine on motor performance predicts the type of response to placebo. *Psychopharmacology*, 106(2), 209-214.

Flaten, M. A. (1998). Information about drug effects modify arousal. An investigation of the placebo response. *Nordic Journal of Psychiatry*, 52(2), 147-151.

- Flaten, M. A., Aasli, O., & Blumenthal, T. D. (2003). Expectations and placebo responses to caffeine-associated stimuli. *Psychopharmacology*, 169(2), 198-204. doi: 10.1007/s00213-003-1497-8
- Flaten, M. A., & Blumenthal, T. D. (1999). Caffeine-associated stimuli elicit conditioned responses: An experimental model of the placebo effect. *Psychopharmacology*, 145(1), 105-112.
- Flaten, M. A., Simonsen, T. C., & Olsen, H. (1999). Drug-Related Information Generates Placebo and Nocebo Responses That Modify the Drug Response. *Psychosomatic Medicine March/April*, 61(2), 250-255.
- Gavrylyuk, G., Ehrt, O., & Meissner, K. (2010). Lack of expectancy-induced placebo effects on pupil size and accommodation. *Zeitschrift fur Medizinische Psychologie*, 19(3-4), 154-160.
- Geers, A. L., Helfer, S. G., Kosbab, K., Weiland, P. E., & Landry, S. J. (2005). Reconsidering the role of personality in placebo effects: Dispositional optimism, situational expectations, and the placebo response. *Journal of Psychosomatic Research*, 58(2), 121-127. doi: 10.1016/j.jpsychores.2004.08.011
- Geers, A. L., Helfer, S. G., Weiland, P. E., & Kosbab, K. (2006). Expectations and placebo response: A laboratory investigation into the role of somatic focus. *Journal of Behavioral Medicine*, 29(2), 171-178. doi: 10.1007/s10865-005-9040-5
- Geers, A. L., Kosbab, K., Helfer, S. G., Weiland, P. E., & Wellman, J. A. (2007). Further evidence for individual differences in placebo responding: An interactionist perspective. *Journal of Psychosomatic Research*, 62(5), 563-570. doi: 10.1016/j.jpsychores.2006.12.005

- Geers, A. L., Weiland, P. E., Kosbab, K., Landry, S. J., & Helfer, S. G. (2005). Goal activation, expectations, and the placebo effect. *Journal of Personality & Social Psychology*, 89(2), 143-159.
- Geers, A. L., Wellman, J. A., Fowler, S. L., Rasinski, H. M., & Helfer, S. G. (2011). Placebo expectations and the detection of somatic information. *Journal of Behavioral Medicine*, 34(3), 208-217. doi: 10.1007/s10865-010-9301-9
- Gibbons, F. X., Carver, C. S., Scheier, M. F., & Hormuth, S. E. (1979). Self-focused attention and the placebo effect: Fooling some of the people some of the time. *Journal of Experimental Social Psychology*, 15(3), 263-274.
- Goetz, C. G., Laska, E., Hicking, C., Damier, P., Muller, T., Nutt, J., . . . Russ, H. (2008). Placebo influences on dyskinesia in Parkinson's disease. *Movement Disorders*, 23(5), 700-707. doi: 10.1002/mds.21897
- Goldman, A. R., Witton, K., & Scherer, J. M. (1965). The drug-giving ritual, verbal instructions and schizophrenics ward activity levels. *Journal of Nervous and Mental Disease*, 140(4), 272-279.
- Hahn, R. (1997). The Nocebo Phenomenon: Concept, Evidence, and Implications for Public Health. *Preventive Medicine*, 26, 607-611.
- Harrell, P. T., & Juliano, L. M. (2009). Caffeine expectancies influence the subjective and behavioral effects of caffeine. *Psychopharmacology*, 207(2), 335-342.
- Harrell, P. T., & Juliano, L. M. (2012). A Direct Test of the Influence of Nicotine Response Expectancies on the Subjective and Cognitive Effects of Smoking. *Experimental and Clinical Psychopharmacology*, 20(4), 278-286. doi: 10.1037/a0028652
- Hauser, W., Bartram, C., Bartram-Wunn, E., & Tolle, T. (2012). Adverse Events Attributable to Nocebo in Randomized Controlled Drug Trials in Fibromyalgia Syndrome and

- Painful Diabetic Peripheral Neuropathy Systematic Review. *Clinical Journal of Pain*, 28(5), 437-451.
- Hauser, W., Hansen, E., & Enck, P. (2012). Nocebo Phenomena in Medicine Their Relevance in Everyday Clinical Practice. *Deutsches Arzteblatt International*, 109(26), 459-466.
- Heatherton, T. F., Polivy, J., & Herman, C. P. (1989). Restraint and Internal Responsiveness: Effects of Placebo Manipulations of Hunger State on Eating. *Journal of Abnormal Psychology*, 98(1), 89-92.
- Higgins, J. P. T., Altman, D. G., Gotzsche, P. C., Juni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343. doi: 10.1136/bmj.d5928
- Higuchi, T., Shoji, K., & Hatayama, T. (2002). Smelling lavender and jasmine with advance information about their psychological effects: An examination of the placebo effect. *Tohoku Psychological Folia*, 61, 1-10.
- Jaen, C., & Dalton, P. (2014). Asthma and odors: The role of risk perception in asthma exacerbation. *Journal of Psychosomatic Research*, 77(4), 302-308. doi: 10.1016/j.jpsychores.2014.07.002
- Jensen, M. P., & Karoly, P. (1991). Motivation and expectancy factors in symptom perception: A laboratory study of the placebo effect. *Psychosomatic Medicine*, 53(2), 144-152.
- Kam-Hansen, S., Jakubowski, M., Kelley, J. M., Kirsch, I., Hoaglin, D. C., Kaptchuk, T. J., & Burstein, R. (2014). Altered Placebo and Drug Labeling Changes the Outcome of Episodic Migraine Attacks. *Science Translational Medicine*, 6(218). doi: 10.1126/scitranslmed.3006175

- Kaptchuk, T. J., Stason, W. B., Davis, R. B., Legedza, A. T., Schnyer, R. N., Kerr, C. E., . . .
- Goldman, R. H. (2006). Sham device v inert pill: Randomised controlled trial of two placebo treatments. *BMJ: British Medical Journal*, 332(7538), 391-397.
- Kennedy, W. P. (1961). The nocebo reaction. *Medicina experimentalis. International journal of experimental medicine*, 95, 203-205.
- Kirsch, I. (1997a). Response expectancy theory and application: A decennial review. *Applied and Preventive Psychology*, 6(2), 69-79.
- Kirsch, I. (1997b). Specifying nonspecifics: Psychological mechanisms of placebo effects. In A. Harrington (Ed.), *The placebo effect: An interdisciplinary exploration* (pp. 166-186). Cambridge, MA: Harvard University Press.
- Kirsch, I., & Weixel, L. J. (1988). Double-Blind Versus Deceptive Administration of a Placebo. *Behavioral Neuroscience*, 102(2), 319-323.
- Köteles, F., & Babulka, P. (2014). Role of expectations and pleasantness of essential oils in their acute effects. *Acta Physiologica Hungarica*, 101(3), 329-340.
- Kuenzel, J., Blanchette, I., Zandstra, E. H., Thomas, A., & El-Deredy, W. (2012). Awareness changes placebo effects for feeling relaxed, but not for liking. *Journal of Marketing Communications*, 18(5), 379-396.
- Liccardi, G., Senna, G., Russo, M., Bonadonna, P., Crivellaro, M., Dama, A., . . .
- Passalacqua, G. (2004). Evaluation of the nocebo effect during oral challenge in patients with adverse drug reactions. *Journal of Investigational Allergology and Clinical Immunology*, 14(2), 104-107.
- Link, J., Haggard, R., Kelly, K., & Forrer, D. (2006). Placebo/nocebo symptom reporting in a sham herbal supplement trial. *Evaluation & the Health Professions*, 29(4), 394-406.
doi: 10.1177/0163278706293403

- Lombardi, C., Gargioni, S., Canonica, G. W., & Passalacqua, G. (2008). The nocebo effect during oral challenge in subjects with adverse drug reactions. *European Annals of Allergy and Clinical Immunology*, 40(4), 138-141.
- Lorber, W., Mazzoni, G., & Kirsch, I. (2007). Illness by suggestion: Expectancy, modeling, and gender in the production of psychosomatic symptoms. *Annals of Behavioral Medicine*, 33(1), 112-116.
- Lotshaw, S. C., Bradley, J. R., & Brooks, L. R. (1996). Illustrating caffeine's pharmacological and expectancy effects utilizing a balanced placebo design. *Journal of Drug Education*, 26(1), 13-24.
- Mazzoni, G., Foan, L., Hyland, M. E., & Kirsch, I. (2010). The Effects of Observation and Gender on Psychogenic Symptoms. *Health Psychology*, 29(2), 181-185. doi: 10.1037/a0017860
- Meulders, A., Fannes, S., Van Diest, I., De Peuter, S., Vansteenwegen, D., & Van den Bergh, O. (2010). Resistance to extinction in an odor–20% CO₂ inhalation paradigm: Further evidence for a symptom learning account of multiple chemical sensitivity. *Journal of Psychosomatic Research*, 68(1), 47-56. doi: <http://dx.doi.org/10.1016/j.jpsychores.2009.03.009>
- Mikalsen, A., Bertelsen, B., & Flaten, M. A. (2001). Effects of caffeine, caffeine-associated stimuli, and caffeine-related information on physiological and psychological arousal. *Psychopharmacology*, 157(4), 373-380. doi: 10.1007/s002130100841
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Plos Medicine*, 6(7). doi: 10.1371/journal.pmed.1000097
- Molcan, J., & et al. (1982). Perceived effect of placebo unrelated to the effect imputed to a fancied drug. *Activitas Nervosa Superior*, 24(4), 270-271.

- Monguio, R. R., Otero, M. J., & Rovira, J. (2003). Assessing the Economic Impact of Adverse Drug Effects. *Pharmacoeconomics*, 21(9), 623-650.
- Mrna, B., & Skrivanek, A. (1985). Placebo effect on healthy volunteers-athletes. *Activitas Nervosa Superior*, 27(1), 42-43.
- Myers, M. G., Cairns, J. A., & Singer, J. S. (1987). The consent form as a possible cause of side effects. *Clinical Pharmacological Therapy*, 42, 250-253.
- Neukirch, N., & Colagiuri, B. (2014). The placebo effect, sleep difficulty, and side effects: a balanced placebo model. *Journal of Behavioral Medicine*.
- Nevelsteen, S., Legros, J. J., & Crasson, M. (2007). Effects of information and 50 Hz magnetic fields on cognitive performance and reported symptoms. *Bioelectromagnetics*, 28(1), 53-63.
- NICE. (2009). *Medicines adherence: Involving patients in decisions about prescribed medicines and supporting adherence*.
<http://www.nice.org.uk/guidance/cg76/resources/guidance-medicines-adherence-pdf>.
- Ossege, M., Sycha, T., Aigner, M., Schmetterer, L., Eichler, H., Muller, M., . . . Bauer, P. (2005). Effect of information on reported adverse events in a placebo-controlled trial. *Drug Safety*, 28(1), 81-87. doi: 10.2165/00002018-200528010-00006
- Papoiu, A. D. P., Wang, H., Coghill, R. C., Chan, Y. H., & Yosipovitch, G. (2011). Contagious itch in humans: A study of visual 'transmission' of itch in atopic dermatitis and healthy subjects. *British Journal of Dermatology*, 164(6), 1299-1303.
- Penick, S., & Fisher, S. (1965). Drug-set interaction: Psychological and physiological effects of epinephrine under differential expectation. *Psychosomatic Medicine*, 27(2), 177-182.
- Pennebaker, J. W., & Skelton, J. A. (1981). Selective Monitoring of Physical Sensations. *Journal of Personality and Social Psychology*, 41(2), 213-223.

- Petersen, G. J., Finnerup, N. B., Colloca, L., Amanzio, M., Price, D. P., Jensen, T. S., & Vase, L. (2014). The magnitude of nocebo effects in pain: A meta-analysis. *Pain, 155*, 1426-1434.
- Petrie, K. J., Broadbent, E., Kley, N., Moss-Morris, R., Horne, R., & Rief, W. (2005). Worries about modernity predict symptom complaints after environmental pesticide spraying. *Psychosomatic Medicine, 67*, 778-782.
- Petrie, K. J., Faasse, K., Critchon, F., & Grey, A. (2014). How common are symptoms? Evidence from a New Zealand national telephone survey. *BMJ Open, 4*. doi: 10.1136/bmjopen-2014-005374
- Petrie, K. J., Moss-Morris, R., Grey, C., & Shaw, M. (2004). The relationship of negative affect and perceived sensitivity to symptom reporting following vaccination. *British Journal of Health Psychology, 9*, 101-111.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Britten, N., ... Duffy, S. (2006). Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: Final Report. Swindon: ESRC Methods Programme.
- Price, D. D., Milling, L. S., Kirsch, I., Duff, A., Montgomery, G. H., & Nicholls, S. S. (1999). An analysis of factors that contribute to the magnitude of placebo analgesia in an experimental paradigm. *Pain, 83*(2), 147-156.
- Put, C., Van den Bergh, O., Van Ongeval, E., De Peuter, S., Demedts, M., & Verleden, G. (2004). Negative affectivity and the influence of suggestion on asthma symptoms. *Journal of Psychosomatic Research, 57*(3), 249-255.
- Read, J. C. A., & Bohr, I. (2014). User experience while viewing stereoscopic 3D television. *Ergonomics, 57*(8), 1140-1153. doi: 10.1080/00140139.2014.914581
- Rubin, G. J., Cleare, A. J., & Wessely, S. (2008). Psychological factors associated with self-reported sensitivity to mobile phones. *Journal of Psychosomatic Research, 64*(1), 1-9.

- Schneider, R., Gruner, M., Heiland, A., Keller, M., Kujanova, Z., Peper, M., . . . Walach, H. (2006). Effects of expectation and caffeine on arousal, well-being, and reaction time. *International Journal of Behavioral Medicine*, 13(4), 330-339. doi: 10.1207/s15327558ijbm1304_8
- Schweiger, A., & Parducci, A. (1981). Nocebo: The psychologic induction of pain. *The Pavlovian journal of biological science : official journal of the Pavlovian*, 16(3), 140-143.
- Slanska, J., Tikal, K., Hvizdosova, J., & Benesova, O. (1974). Placebo effect related to the type of indoctrination and some personality traits. *Ceskoslovenska Psychiatrie*, 70(3), 174-179.
- Stegen, K., Neujens, A., Crombez, G., Hermans, D., Van de Woestijne, K., & Van den Bergh, O. (1998). Negative affect, respiratory reactivity, and somatic complaints in a CO2-enriched air inhalation paradigm. *Biological Psychology*, 49, 109-122.
- Stegen, K., Van Diest, I., Van de Woestijne, K., & Van den Bergh, O. (2000). Negative affectivity and bodily sensations induced by 5.5% CO2 enriched air inhalation: Is there a bias to interpret bodily sensations negatively in persons with negative affect? *Psychology & Health*, 15(4), 513-525.
- Stewart-Williams, S. (2004). The placebo puzzle: Putting together the pieces. *Health Psychology*, 23, 198-206. doi: 10.1037/0278-6133.23.2.198
- Strohle, A. (2000). Increased response to a putative panicogenic nocebo administration in female patients with panic disorder. *Journal of Psychiatric Research*, 34(6), 439-442. doi: 10.1016/s0022-3956(00)00039-x
- Suchman, A. L., & Ader, R. (1992). Classic conditioning and placebo effects in crossover studies. *Clinical Pharmacology and Therapeutics*, 52(4), 372-377.

- Sullivan, M. J. L., Lynch, M. E., Clark, A. J., Mankovsky, T., & Sawynok, J. (2008). Catastrophizing and treatment outcome: Differential impact on response to placebo and active treatment outcome. *Contemporary Hypnosis*, 25(3-4), 129-140.
- Symon, A., Williams, B., Adelasoye, Q. A., & Cheyne, H. (2015). Nocebo and the potential harm of 'high risk' labelling: a scoping review. *Journal of Advanced Nursing*. doi: 10.1111/jan.12637
- Szemerszky, R., Dömötör, Z., Berkes, T., & Koteles, F. (2016). Attribution-Based Nocebo Effects. Perceived Effects of a Placebo Pill and a Sham Magnetic Field on Cognitive Performance and Somatic Symptoms. *International Journal of Behavioural Medicine*, 23, 204-213. doi: 10.1007/s12529-015-9511-1
- Szemerszky, R., Koteles, F., Lihi, R., & Bardos, G. (2010). Polluted places or polluted minds? An experimental sham-exposure study on background psychological factors of symptom formation in 'Idiopathic Environmental Intolerance attributed to electromagnetic fields'. *International Journal of Hygiene and Environmental Health*, 213(5), 387-394. doi: 10.1016/j.ijheh.2010.05.001
- Tippens, K. M., Purnell, J. Q., Gregory, W. L., Connelly, E., Hanes, D., Oken, B., & Calabrese, C. (2014). Expectancy, Self-Efficacy, and Placebo Effect of a Sham Supplement for Weight Loss in Obese Adults. *Journal of Evidence-Based Complementary and Alternative Medicine*, 19(3), 181-188.
- Van den Bergh, O., Kempynck, P. J., van de Woestijne, K. P., Baeyens, F., & Eelen, P. (1995). Respiratory learning and somatic complaints: a conditioning approach using CO₂-enriched air inhalation. *Behaviour Research & Therapy*, 33(5), 517-527.
- Van den Bergh, O., Stegen, K., & Van de Woestijne, K. P. (1997). Learning to have psychosomatic complaints: conditioning of respiratory behavior and somatic complaints in psychosomatic patients. *Psychosomatic Medicine*, 59(1), 13-23.

- Van den Bergh, O., Stegen, K., & Van de Woestijne, K. P. (1998). Memory effects on symptom reporting in a respiratory learning paradigm. *Health Psychology, 17*(3), 241-248.
- Van den Bergh, O., Stegen, K., Van Diest, I., Raes, C., Stulens, P., Eelen, P., . . . Nemery, B. (1999). Acquisition and extinction of somatic symptoms in response to odours: a Pavlovian paradigm relevant to multiple chemical sensitivity. *Occupational & Environmental Medicine, 56*(5), 295-301.
- Van Diest, I., De Peuter, S., Piedfort, K., Bresseleers, J., Devriese, S., Van de Woestijne, K. P., & Van den Bergh, O. (2006). Acquired lightheadedness in response to odors after hyperventilation. *Psychosomatic Medicine, 68*(2), 340-347.
- Vase, L. a. b., Baram, S. c. d., Takakura, N. e., Yajima, H. e., Takayama, M. e., Kaptchuk, T. J. f., . . . Svensson, P. c. (2013). Specifying the nonspecific components of acupuncture analgesia. *Pain, 154*(9), 1659-1667.
- Vits, S., Cesko, E., Benson, S., Rueckert, A., Hillen, U., Schadendorf, D., & Schedlowski, M. (2013). Cognitive Factors Mediate Placebo Responses in Patients with House Dust Mite Allergy. *Plos One, 8*(11). doi: 10.1371/journal.pone.0079576
- Vogtle, E., Barke, A., & Kroner-Herwig, B. (2013). Nocebo hyperalgesia induced by social observational learning. *Pain, 154*, 1427-1433.
- Walach, H., Schmidt, S., Bihr, Y. M., & Wiesch, S. (2001). The Effects of a Caffeine Placebo and Experimenter Expectation on Blood Pressure, Heart Rate, Well-Being, and Cognitive Performance. *European Psychologist, 6*(1), 15-25. doi: 10.1027//1016-9040.6.1.15
- Walach, H., Schmidt, S., Dirhold, T., & Nosch, S. (2002). The effects of a caffeine placebo and suggestion on blood pressure, heart rate, well-being and cognitive performance.

International Journal of Psychophysiology, 43(3), 247-260. doi: 10.1016/s0167-8760(01)00188-x

Walach, H., & Schneider, R. (2009). Does the presence of a pharmacological substance alter the placebo effect?-results of two experimental studies using the placebo-caffeine paradigm. *Human Psychopharmacology-Clinical and Experimental*, 24(7), 549-558. doi: 10.1002/hup.1054

Wells, R. E., & Kaptchuk, T. J. (2012). To Tell the Truth, the Whole Truth, May Do Patients Harm: The Problem of the Nocebo Effect for Informed Consent. *The American Journal of Bioethics*, 12(3), 22-29.

Wendt, L., Albring, A., Benson, S., Engler, H., Engler, A., Hinney, A., . . . Schedlowski, M. (2014). Catechol-O-Methyltransferase Val158Met Polymorphism Is Associated with Somatosensory Amplification and Nocebo Responses. *Plos One*, 9(9). doi: 10.1371/journal.pone.0107665

Winters, W., Devriese, S., Eelen, P., Veulemans, H., Nemery, B., & Van den Bergh, O. (2001). Symptom Learning in Response to Odors in a Single Odor Respiratory Learning Paradigm. *Annals of the New York Academy of Sciences*, 933(1), 315-318. doi: 10.1111/j.1749-6632.2001.tb05834.x

Winters, W., Devriese, S., Van Diest, I., Nemery, B., Veulemans, H., Eelen, P., . . . Van den Bergh, O. (2003). Media warnings about environmental pollution facilitate the acquisition of symptoms in response to chemical substances. *Psychosomatic Medicine*, 65(3), 332-338.

Wise, R. A., Bartlett, S. J., Brown, E. D., Castro, M., Cohen, R., Holbrook, J. T., . . . American Lung Association Asthma Clinical Research, C. (2009). Randomized trial of the effect of drug presentation on asthma outcomes: the American Lung

Association Asthma Clinical Research Centers. *Journal of Allergy & Clinical Immunology*, 124(3), 436-444.

Witthöft, M., & Rubin, G. J. (2013). Are media warnings about the adverse health effects of modern life self-fulfilling? An experimental study on idiopathic environmental intolerance attributed to electromagnetic fields (IEI-EMF). *Journal of Psychosomatic Research*, 74(3), 206-212.

Zimmermann-Viehoff, F., Meissner, K., Koch, J., Weber, C. S., Richter, S., & Deter, H. C. (2013). Autonomic effects of suggestive placebo interventions to increase or decrease blood pressure: A randomized controlled trial in healthy subjects. *Journal of Psychosomatic Research*, 75(1), 32-35. doi: 10.1016/j.jpsychores.2013.03.011

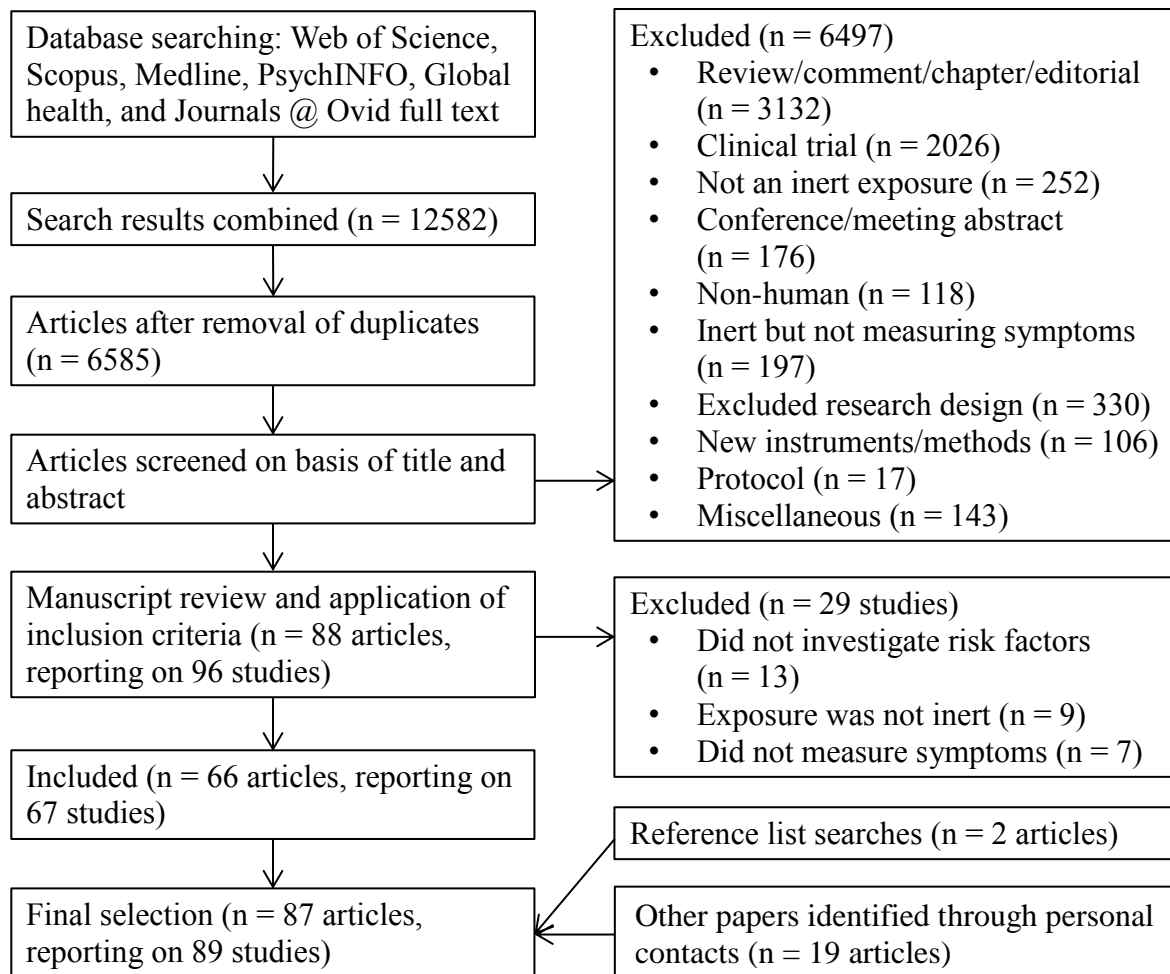
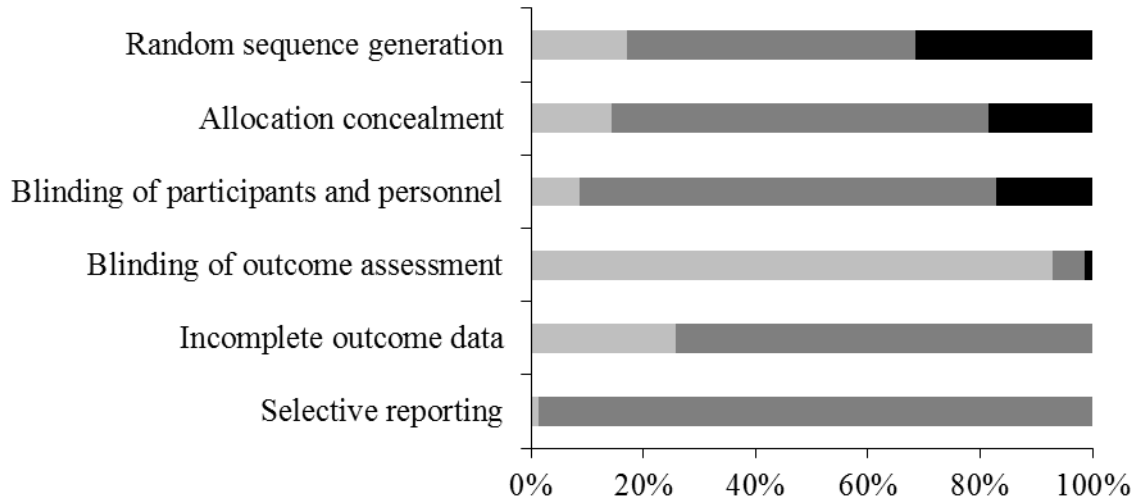


Figure 1. Flow diagram of the selection process of studies including the number of events and reasons for exclusion.

a) Experimental studies



b) Prospective studies

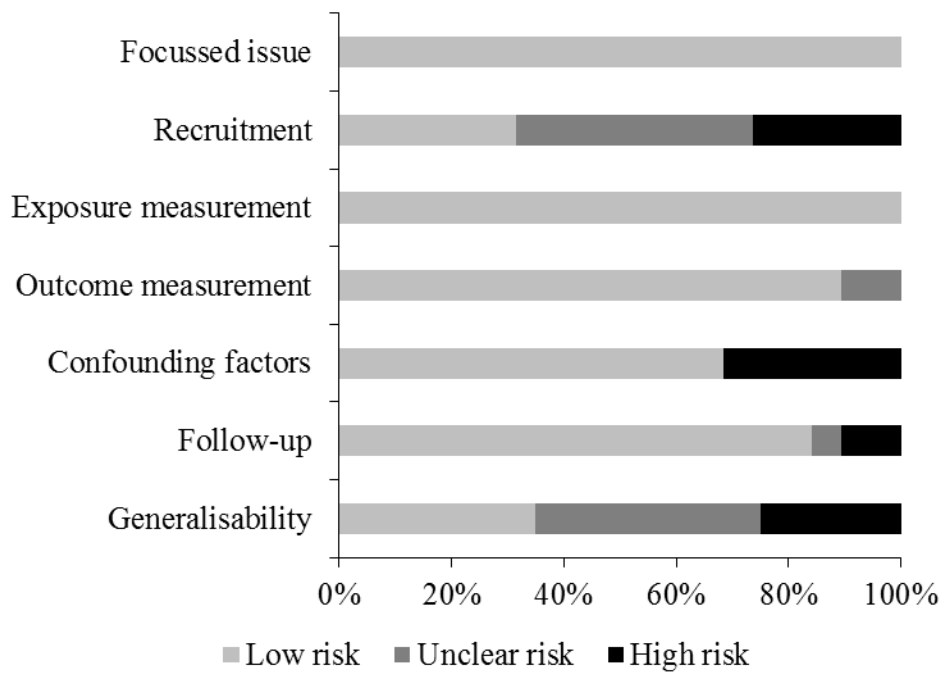


Figure 2. Quality assessment of experimental and prospective studies

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Table 1. Summary of the methods used in the experimental studies

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental risk factor(s) and conditions	Baseline risk factors
Andre-Obadia et al. (2011) b,e	RCT (B)	Chronic neuropathic pain patients (45, 55.0, 37.8)	Sham rTMS	1. Prior experience: a. Sham rTMS before active rTMS (20); b. Sham rTMS after successful active rTMS (12); c. Sham rTMS after ineffective active rTMS (13)	Pain
Angelucci and Pena (1997) e	RCT (B)	Student caffeine consumers (148, U/K, 23.0)	Sham coffee	1. Arousal suggestions: a. Given coffee with no expectations (37); b. Given coffee with low arousal expectations (37); c. Given coffee with high arousal expectations (37); d. no coffee and no expectations (37)	State and trait anxiety, Suggestibility, Expectations, Gender
Bayer et al. (1991) e	RCT (B+W)	Unemployed Men (100, U/K, 100.0)	Sham electrical shock	1. Symptom suggestions: a. Told they would receive a safe but often painful undetectable current (60); b. Were assured there would be no shocks (40) 2. Perceived dose: a. Within each group the stimulator setting increased from 0 to 80 mA	None
Bayer et al. (1998) a,e	RCT (B+W)	Job seekers (62, U/K, 82.0)	Sham electrical shock	1. Prior experience: a. Exposed to two physical pain induction procedures prior to sham stimulation (32); b. Warned of pain and received sham stimulation. They were not exposed to any prior pain induction (30) 2. Perceived dose: a. Within each group the stimulator setting increased in steps of 10 every 5 minutes till it reached 50	Expectations
Benedetti et al. (1997) e	RCT (B)	Video assisted thoracoscopy patients (36, 53.7, 66.1)	Sham treatment	1. Symptom suggestions: a. Open injection that it would increase pain (18); b. Hidden injection (18)	None
Brodeur (1965)	RCT (B)	Healthy senior students	Sham arousal capsule	1. Arousal suggestions: a. Told it was a stimulant (15); b. Told it was a tranquilizer (15); c. No suggestion (15)	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

e		(45, U/K, 91.1)			
Colagiuri et al. (2012)	RCT (B)	Students experiencing sleep difficulty (82, 20.2, 22.0)	Sham sleeping pill	1. Symptom suggestions: a. Treatment might cause one side effect (29); b. Treatment might cause four side effects (23); c. No warning about side effects (30)	None
e					
Crichton et al. (2014)	RCT (B)	Students (54, U/K, 37.0)	Sham infrasound	1. Symptom suggestions: a. TV footage detailing symptomatic experiences attributed to wind farms (27); b. TV footage with experts stating wind farms would not cause symptoms (27)	None
e					
Dalton (1999)	RCT (B)	Healthy volunteers (180, 31.7, 49.4)	Odours	1. Odours: a. Pleasant smelling methyl salicylate (60); b. neutral smelling isobornyl acetate (60); c. Foul smelling butanol (60) 2. Symptom suggestions: a. Told they would have relaxing effects (60); b. Told they were industrial solvents (60); c. Told they were approved for olfactory research (60)	Odour reactivity, Olfactory sensitivity
e					
De Peuter et al. (2005)	RCT (W)	Asthma patients and healthy controls (40, 23.9, 52.5)	Sham inhaler	1. Conditioning: a. one sham inhaler paired with CO2 challenge; b. one sham inhaler paired with O2	Expectations, Negative affect, Clinical condition
e					
Devriese et al. (2000)	Non RCT (B+W)	Healthy students (56, U/K, 41.1)	Odours	1. Odour: a. Foul smelling ammonia; b. Pleasant smelling niaouli 2. Conditioning: a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (28); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (28) 3. Timing: a. Test phase immediately after conditioning trials (28); b. Test phase one week after conditioning trials (28) 4. Generalisation: a. New foul smelling odour butyric acid; b. New foul smelling odour acetic acid; c. New	Negative affect

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				pleasant smelling odour citric aroma	
Devriese et al. (2004) a,e	Non RCT (B+W)	Healthy students (53, U/K, U/K)	Odours	1. Odour: a. Foul smelling ammonia; b. Foul smelling butyric acid 2. Conditioning: a. Ammonia paired with CO2 breathing task, butyric acid paired with room air breathing task (28); b. Ammonia paired with room air breathing task, butyric acid paired with CO2 breathing task (25) 3. Symptom suggestions: a. Given information about possible health damaging effects of chemical pollution (U/K); b. No information (U/K)	Negative affect, Perceived cue odour
Devriese et al. (2006)	RCT (B+W)	Psychology students (40, U/K, 0.0)	Odours	1. Odour: a. Foul smelling ammonia; b. Foul smelling acetic acid 2. Conditioning : a. Ammonia paired with CO2 breathing task, acetic acid paired with room air breathing task (20); b. Ammonia paired with room air breathing task, acetic acid paired with CO2 breathing task (20) 3. Symptom suggestions: a. Given information about possible health damaging effects of chemical pollution (20); b. No information (20)	None
Dinnerstein and Halm (1970) c,e	RCT (B)	Male students (80, U/K, 100.0)	Sham arousal liquid	1. Arousal suggestions: a. Told it was an energizer (40); b. Told it was a tranquilizer (40) 2. Prior experience: a. Received aspirin prior to sham (40); b. Received lactose prior to sham (40)	None
Faasse et al. (2013) b,c,e	RCT (B)	Healthy students (60, 19.4, 43.5)	Sham anti-anxiety tablet	1. Brand suggestions: a. Branded reformulation change (20); b. Generic reformulation change (20); c. No change (20)	None
Flaten (1998) e	RCT (B)	Healthy students (48, U/K, 35.4)	Sham arousal drink	1. Arousal suggestions: a. Told you will feel relaxed and sleepy (16); b. Told you will feel alert and a little stress (16); c. Told you will take an inactive drug (16)	None
Flaten and Blumenthal	RCT (W)	Healthy coffee drinkers	Decaffeinated solution	1. Association: a. Orange juice; b. Decaffeinated coffee	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

(1999) e		(21, 24.8, 61.9)			
Flaten et al. (1999) e	RCT (B)	Healthy volunteers in non-health professions (34, U/K, 54.5)	Sham arousal capsule	1. Arousal suggestions: a. The drug will make you feel relaxed (11); b. The drug will make you feel alert (12); c. You will receive capsules that contain a prescription drug (11)	None
Flaten et al. (2003) a,b,e	W	Coffee drinkers (20, U/K, 50.0)	Sham coffee	1. Perceived dose: a. Participants were first given one cup and then a second	Symptoms, Expectations
Gavrylyuk et al. (2010) e	RCT (B)	Healthy volunteers (30, 24.9, 32.0)	Saline eye drops	1. Symptom suggestions: a. Informed of pupil dilation effects (10); b. Informed of pupil constriction effects (10); c. Informed of saline eye drops (10)	None
Geers et al. (2006) e	RCT (B)	Healthy students (54, U/K, 31.5)	Sham over-the-counter pill	1. Likelihood suggestions: a. Told the pill had unpleasant side effects (18); b. Told they may or may not receive the active drug (19); c. Told they would ingest an inactive drug (17) 2. Self-awareness: a. Told to closely monitor feelings/bodily sensations (27); b. Not given any such instructions (27)	None
Geers et al. (2011) e	RCT (B)	Healthy students (102, 20.5, 21.6)	Sham caffeine capsule	1. Likelihood suggestions: a. Told it contained 250mg of caffeine (34); b. Told they may or may not be ingesting 250mg of caffeine (34); c. Not given the capsule and received no caffeine expectation (34)	Gender, Age, Caffeine consumptions
Geers, Helfer, et al. (2005) e	RCT (B)	Healthy students (54, 21.0, 29.6)	Sham over-the-counter pill	1. Likelihood suggestions: a. Told the pill had unpleasant side effects (18); b. Told the pill would make them feel either unpleasant or was an inactive substance (18); c. Told they would ingest an inactive pill (18) 2. Self-awareness: a. Told to attend to any symptoms experienced (27); b. Not given any such instructions (27)	Age, Gender, Optimism
Geers,	RCT	Healthy students	Sham	1. Arousal suggestions: a. Told they were given caffeine	Caffeine consumption

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Weiland, et al. (2005) e	(B)	(57, U/K, 35.1)	caffeine pill	(U/K); b. No mention of caffeine (U/K) 2. Cooperation prime: a. Given a scrambled sentence test with a cooperation prime (U/K); b. Given a scrambled sentence test with a neutral prime (U/K)	
Gibbons et al. (1979) a,e	RCT (B)	Female students (38, U/K, 0.0)	Sham drug	1. Symptom suggestions: a. Told they were taking Cavanol which would produce some noticeable side effects (19); b. Told they were taking baking soda (19) 2. Self-awareness: a. Mirror was facing participants (19); b. Mirror was not facing participants (19)	None
Goldman et al. (1965) a,b,e	Non RCT (B)	Male veterans with schizophrenia (64, 44.0, 100.0)	Sham arousal treatment	1. Type of administration: a. Received sugar pill (32); b. Received saline injection (32) 2. Arousal suggestoins: a. Told it would heighten their ward activity (32); b. Told it would lower their ward activity (32)	Attitudes towards medication
Harrell and Juliano (2009) c	RCT (B)	Adult non-smoking coffee consumers (30, 22.6, 22.0)	Sham coffee	1. Performance suggestions: a. Told caffeine enhances performance (15); b. Told caffeine impairs performance (15)	None
Harrell and Juliano (2012) c,e	RCT (B)	Adult smokers (43, 28.7, 67.4)	Sham cigarette	1. Performance suggestions: a. Told cigarette enhances performance (20); b. Told cigarette impairs performance (23)	Gender
Heatherton et al. (1989) e	RCT (B)	Female students (59, U/K, 0.0)	Sham vitamin pill	1. Symptom suggestions: a. Told vitamin has been reported to make people feel hungry (19); b. Told vitamin has been reported to make people feel full (20); c. Told no further information (20)	Participant restraint
Higuchi et al. (2002) e	RCT (B)	Healthy volunteers (30, 21.2, 40.0)	Fragrance (Jasmine or Lavendar)	1. Arousal suggestions: a. Told it was relaxing (10); b. Told it was stimulating (10); c. No information given (10)	None
Jaen and Dalton (2014) a,b,e	Non RCT (B)	Asthmatics (17, 38.5, 52.9)	Sham active odour	1. Symptom suggestions: a. Labelled the odour as therapeutic (9); b. Labelled the odour as asthmogenic (8)	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Jensen and Karoly (1991) e	RCT (B+W)	Students (86, U/K, 45.3)	Sham sedative pill	1. Social desirability: a. Type B personality is more positive than type A. Type B have been shown to respond more to pills (43); b. Relationship between type A and B personality and response to pills is very weak (43) 2. Perceived dose: a. Suggestions of a high dose or low dose were counterbalanced across each group	Gender
Kaptchuk et al. (2006)	RCT (B)	Adults with distal pain in the arms (266, 36.7, 45.9)	Sham treatment	1. Type of administration: a. Received sham acupuncture (133); b. Received placebo pill (133)	None
Kirsch and Weixel (1988) e	RCT (B)	Student coffee drinkers (U/K, 19.3, 31.0)	Sham coffee	1. Likelihood suggestions: a. Told they would receive coffee (U/K); b. Told they may or may not receive caffeinated coffee (U/K); c. No beverage, waited for 20 minutes (U/K) 2. Perceived dose: a. 1 tsp (U/K); b. 2 tsps (U/K); c. 3 tsps (U/K); d. 5 tsps (U/K); e. 8 tsps (U/K)	None
Kuenzel et al. (2012) e	RCT (B)	English speaking students (148, 21.7, 18.2)	Herbal infusion tea	1. Arousal suggestions: a. Told it would make them feel relaxed (45); b. Told it would make them feel active (53); c. No information given (50)	None
Lorber et al. (2007) e	RCT (B)	Students without upper respiratory conditions (86, U/K, 40.7)	Sham environmental toxin	1. Social observation: a. Told inhaled substance has been reported to produce symptoms and observed a female confederate inhale and display symptoms (U/K); b. As above but no observation of confederate (U/K); c. Did not inhale the substance and observed a female confederate inhale and display symptoms (U/K); d. As above but no observation of confederate (U/K)	Gender
Lotshaw et al. (1996) e	RCT (B)	Male student coffee drinkers (50, U/K, 100.0)	Sham coffee	1. Arousal suggestions: a. Told coffee received decaffeinated (25); b. Told decaffeinated received decaffeinated (25)	None
Mazzoni et al.	RCT	Healthy students	Sham	1. Social observation: a. Observed a male/female	Personality, Gender,

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

(2010) e	(B)	(120, 20.7, 50.0)	environmental toxin	confederate inhale the substance and display symptoms (60); b. Did not observe a male or female confederate inhale the substance and display symptoms (60)	Gender of model
Meulders et al. (2010) a,e	Non RCT (B+W)	Healthy adults (58, 22.0, 48.3)	Odours	1. Odour: a. Foul smelling ammonia; b. Foul smelling butyric acid 2. Conditioning: a. Ammonia paired with CO2 breathing task, butyric acid paired with room air breathing task (29); b. Ammonia paired with room air breathing task, butyric acid paired with CO2 breathing task (29)	Ability to predict which odour produced the most symptoms
Mikalsen et al. (2001) e	RCT (W)	Student coffee drinkers (21, 25.9, 66.7)	Sham coffee	1. Arousal suggestions: a. Told it was caffeine; b. Told it was not caffeine 2. Association: a. Given in a juice solution; b. Given in a coffee solution	None
Mrna and Skrivaneck (1985) a,b,e	W	Healthy volunteers (21, 17.0, 47.6)	Sham arousal drug	1. Arousal suggestions: a. Told it was a new doping drug undetectable by anti-doping tests; b. Told it was to relax pre-restart states	Prior placebo response
Neukirch and Colagiuri (2014) a,e	RCT (B)	Students with sleep difficulty (91, 21.3, 33.0)	Sham sleep medication	1. Symptom suggestions: a. Warned about an increase/decrease in appetite and received placebo treatment (24); b. Warned about the side effect but received no treatment (23); c. Not warned about the side effects and received placebo treatment (22); d. Not warned about the side effects and received no treatment (22)	None
Nevelsteen et al. (2007) e	RCT (B)	Healthy males (59, 48.4, 100.0)	Sham magnetic field	1. Performance suggestions: a. Told magnetic fields enhance cognitive performance (15); b. Told magnetic fields impair cognitive performance (15); c. Told magnetic fields have no effect on cognitive performance (14); d. Not exposed to sham magnetic field and received no information (15)	State-trait anxiety, Depression, Positive and Negative affect, Sensitivity to anxiety, Vigilance, Comfort under helmet

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Ossege et al. (2005)	RCT (B)	Healthy volunteers (60, 27.6, 40.0)	Placebo drug	1. Likelihood suggestions: a. Misleading information that is was an active medication (30); b. 50% chance that it was a placebo or active medication (30)	None
Papoiu et al. (2011) e	RCT (W)	Healthy volunteers and patients with atopic dermatitis (25, U/K, 44.0)	Sham histamine	1. Social observation: a. Watched a 5 minute video of people scratching their left forearm; b. Watched a 5 minute video of the same persons in the scratching video but sitting idle.	Gender
Penick and Fisher (1965) a,b,c,e	W	Healthy medical students (14, U/K, U/K)	Sham arousal drug	1. Arousal suggestions: a. Told they would receive a stimulant drug; b. Told they would receive a sedative drug	None
Pennebaker and Skelton (1981) e	RCT (B)	Students (38, U/K, 31.6)	Ultrasonic noise	1. Symptom suggestions: a. Told it would increase skin temperature (13); b. Told it would decrease skin temperature (12); c. Told it would have no effect on skin temperature (13)	None
Put et al. (2004) a,b,c,e	W	Asthma patients (32, 40.0, 50.0)	Sham inhaler	1. Symptom suggestions: a. Told it would have no effect on breathing; b. Told it was a bronchoconstrictor; c. Told it was a bronchodilator	Negative affect, Social desirability
Read and Bohr (2014) a,b,c,e	Non RCT (B)	Volunteers without photosensitive epilepsy (177, 25.3, U/K)	Sham 3D TV	1. Symptom suggestions: a. Told it was 3D and wore passive 3D glasses (22); b. Told it was 3D and wore active no shuttering 3D glasses (33); c. Told it was 2D and did not wear glasses (122)	Gender
Schneider et al. (2006) c,e	RCT (B)	Healthy Adults (45, 31.0, 22.2)	Sham coffee	1. Arousal suggestions: a. Told they were to consume decaffeinated coffee (15); b. Told they were to consume regular coffee (15); c. Informed they would receive no beverage and no instructions (15)	None
Schweiger and Parducci (1981) e	RCT (B)	Students (34, U/K, 52.9)	Sham electric current	1. Symptom suggestions: a. Told a low current would be delivered, too mild to be felt but had produced mild headaches in the past (17); b. Told current would be too weak to be felt, but some people develop mild headaches	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				as a side effect (17)	
Slanska et al. (1974) a,e	Non RCT (B)	Medical students (33, U/K, U/K)	Salt solution	1. Arousal suggestions: a. Told it was a stimulant (17); b. Told it was a sedative (16)	Stability – instability, Activity – passivity, Submissive- dominance, Rationality- sensuousness, Introversion- extraversion
Stegen et al. (1998) e	RCT (W)	Healthy psychology students (72, U/K, 48.6)	Breathing trial with room air	1. Conditioning: a. Room air breathing trial before 7.5% CO2 challenge; b. Room air breathing trial after 7.5% CO2 challenge	Negative affect
Szemerszky et al. (2010) a,b,c,e	W	Healthy students (40, 22.8, 27.5)	Sham EMF	1. Perceived dose: a. Told it would be weak; b. Told it would be strong	Gender, Expectations, IEI-EMF scores, State anxiety, Dispositional optimism, Somatisation, Somatosensory amplification, Motivation
Tippens et al. (2014) e	RCT (B)	Obese adults (79, 49.4, 10.4)	Sham weight loss supple- ment	1. Likelihood suggestions: a. Told they would be given an active weight loss supplement (27); b. Told they would be randomly assigned to either the active or placebo supplement (28); c. Only received lifestyle education (24)	None
Van den Bergh et al. (1999) a,e	Non RCT (B+W)	Healthy students (64, U/K, 25.0)	Odours	1. Odour: a. Foul smelling ammonia; b. Foul smelling butyric acid 2. Conditioning: a. Ammonia paired with CO2 breathing task, butyric acid paired with room air breathing task (32); b. Ammonia paired with room air breathing task,	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

butyric acid paired with CO2 breathing task (32)					
Van den Bergh et al. (1995) a,e	Non RCT (B+W)	Healthy students (28, U/K, 50.0)	Odours	1. Odour: a. Foul smelling ammonia; b. Pleasant smelling niaouli 2. Conditioning : a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (14); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (14)	Negative affect
Van den Bergh et al. (1997) a,e	Non RCT (B+W)	Psychosomatic patients (28, 36.0, 50.0)	Odours	1. Odour: a. Foul smelling ammonia; b. Pleasant smelling niaouli 2. Conditioning : a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (14); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (14) 3. Generalisation: a. New foul smelling odour Ichtyol; b. New pleasant smelling odour Rose	Gender, State and trait anxiety, Blunting behaviour
Van den Bergh et al. (1998) e	RCT (B+W)	Healthy adults (56, 42.5, 50.0)	Odours	1. Odour: a. Foul smelling ammonia; b. Pleasant smelling niaouli 2. Self-awareness: a. Told to count lower tones and disregard higher tones (28); b. Told to ignore tones (28) 3. Conditioning: a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (28); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (28) 4. Generalisation: a. New foul smelling odour Ichtyol; b. New pleasant smelling odour Rose	Gender
Van Diest et al. (2006) e	RCT (B+W)	Students (28, U/K, 21.4)	Odours	1. Odour: a. Foul smelling ammonia; b. Foul smelling acetic acid 2. Conditioning : a. Ammonia paired with hypocapnic over breathing trial, acetic acid paired with normocapnic over breathing trial (13); b. Ammonia paired with normocapnic over breathing trial, acetic acid paired with	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				hypocapnic over breathing trial (15) 3. Type of breathing: a. Test odours given with normocapnic breathing trial (U/K); b. Test odours given with spontaneous breathing (U/K)	
Walach and Schneider (2009) Exp 1	RCT (B)	Healthy adult coffee drinkers (60, 32.3, 23.3)	Sham coffee	1. Likelihood suggestions: a. Told it was caffeine (15); b. Told it could be placebo or caffeine (15); c. Told it could be placebo or caffeine (15); d. Received no beverage (15)	Expectations
Walach and Schneider (2009) Exp 2	RCT (B)	Healthy adults coffee drinkers (30, 29.9, 33.3)	Sham coffee	1. Arousal suggestions: a. Told it was caffeine (15); b. Received no beverage (15)	Expectations
Walach et al. (2001)	RCT (B)	Coffee drinkers (157, 28.1, 34.0)	Sham coffee	1. Likelihood suggestions: a. Told they would receive a placebo (41); b. Told they would receive coffee (39); c. Told they may receive real coffee or decaffeinated coffee (39); d. No substance or instruction given (38) 2. Experimenter expectations: a. Experimenter told the physiological effects from a caffeine placebo are real (proplacebo) (U/K); b. Experimenter told the effects of caffeine placebos are just due to artefacts (antiplacebo) (U/K)	Expectations
Walach et al. (2002)	RCT (B)	Coffee drinkers (159, 25.5, 58.0)	Sham coffee	1. Symptom suggestions: a. Received an information leaflet describing the pharmacological effects of caffeine (U/K); b. Received no further information (U/K) 2. Likelihood suggestions: a. Told they would receive a placebo (39); b. Told they would receive coffee (40); c. Told they may receive real coffee or decaffeinated coffee (40); d. No substance or instruction given (40)	None
Winters et al. (2001) Exp 1 a,e	Non RCT (B)	Psychology students (50, U/K,U/K)	Ammonia	1. Conditioning : a. Odour + CO2 trials and room air trials (10); b. Odour trials and CO2 trials (10); c. Odour trials, CO2 trials, odour + CO2 trials, room air trials (10); d. odour trials, room air trials (10); e. CO2 trials, room air trials (10)	None

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Winters et al. (2001) Exp 2 a,e	Non RCT (B)	18-30 year olds (40, U/K,U/K)	Odours	1. Odour: a. Foul smelling ammonia (20); b. Pleasant smelling niaouli (20) 2. Conditioning : a. Odour + CO2 trials and room air trials (20); b. Odour trials and CO2 trials (20)	None
Winters et al. (2003) e	Non RCT (B+W)	18-30 year olds (32, U/K,15.6)	Odours	1. Odour: a. Foul smelling ammonia; b. Pleasant smelling niaouli 2. Conditioning : a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (16); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (16) 3. Verbal suggestions of symptoms: a. Given leaflet describing widespread chemical pollution of the environment is a potential cause of multiple chemical sensitivity (16); b. No information given (16)	None
Wise et al. (2009) c	RCT (B)	Patients with poor asthma control (241, 39.0, 29.5)	Sham asthma drug	a. Emphasized benefit of treatment and described potential side effects (121) b. Expressed uncertainty about improvement following treatment and did not describe potential side effects (120)	None
Witthöft and Rubin (2013)	RCT (B)	Adult English speakers (147, 29.8, 32.7)	Sham EMF	1. Symptom suggestions: a. Watched a documentary concerning the potential adverse health effects of Wi-Fi (76); b. Watched a BBC News report concerning the security of the internet and mobile phone data (71)	State anxiety, Age, Gender, Level of education, Personality
Zimmermann-Viehoff et al. (2013) b,e	RCT (B)	Healthy caucasians (92, 24.5, 41.3)	Sham arousal oral spray	1. Symptom suggestions: a. Told it contained a drug to increase BP (33); b. Told it contained a drug to decrease BP (29); c. Told it was a placebo (30)	None

Note: RCT = Randomised Controlled Trial, Non RCT = Non randomised controlled trial, B = Between subjects design, W = Within subjects design, U/K= Unknown, *italicised* = not directly given but has been extrapolated from the available data, rTMS = Repetitive Transcranial Magnetic Stimulation, EMF = Electromagnetic Field, tsp = Teaspoon, IEI-EMF = Idiopathic environmental intolerance attributed to electromagnetic fields, CO2 = Carbon dioxide, O2 = Oxygen, ns = non-significant, **a** = high risk random sequence generation bias, **b** = high risk

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

allocation concealment bias, **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Table 2. Summary of the methods used in prospective studies

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Baseline risk factor(s)
Bogaerts et al. (2010) e	P	Female patients with medically unexplained dyspnea and healthy controls (58, U/K, 0.0)	Breathing trial with room air	State anxiety, Negative affect, Clinical condition
Casper et al. (2001) e	P	Nonpsychotic major depressive patients (876, U/K, 42.8)	Sham fluoxetine treatment	Gender, Depression severity
Danker-Hopfe et al. (2010)	P	Villages in Germany with weak RF-EMF sources (397, U/K, 49.1)	Sham EMF	Bad sleep quality, General fear/anxiety towards risks of RF-EMF, Fear/anxiety towards base station, Preoccupation with EMF, Visibility of the base station
Davis et al. (1995) a,d,e	P	Healthy Adults (27, U/K, 55.6)	Sham anti-depressant pill	Neuroticism, Somatosensory amplification
de la Cruz et al. (2010) e	P	Patients with cancer related fatigue (105, U/K, 40.0)	Sham treatment	Anxiety, Nausea, Sleep, General health, Well-being, Cognitive status, Age, Education level
De Peuter et al. (2007) e	P	Asthma patients (30, 38.0, 26.7)	Sham histamine inhalation	Negative affect
Drici et al. (1995) b,e	P	Healthy volunteers (52, 23.5, 50.0)	Sham paracetamol eye drop	Employment, Type A Personality, Type B Personality
Fillmore and Vogel-Sprott (1992) e	P	Male students (56, U/K, 100.0)	Sham coffee	Symptom expectations
Goetz et al. (2008) e	P	Parkinson's patients with dyskinesia (484, U/K, U/K)	Sham medication	Age, Gender, Dyskinesia severity, UPDRS motor score, Daily L-dopa dose, Dyskinesia duration, Adverse events, Severity of adverse events,

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				Geographical site of enrolment, Study (1 or 2)
Köteles and Babulka (2014) a,d,e	P	Adult volunteers (33, 37.7, 15.2)	3 types of Essential oils (Randomised to 1)	Expectations, Pleasantness of odour
Liccardi et al. (2004) b,e	P	Patients with ADRs (600, 42.0, 30.3)	Sham allergen pill	Gender, Hospital centre
Link et al. (2006) a,b,c,d,e	P	Students (36, 22.7, 44.0)	Sham herbal supplement	Expectations, State anxiety, Social desirability
Lombardi et al. (2008) a,d,e	P	Patients with ADRs (435, 39.7, 32.0)	Sham allergen pill	Gender, Age, Atopic status, Severity of previous reaction, Type of previous reaction
Molcan et al. (1982) b,e	P	Medical students (48, U/K, 52.1)	Sham arousal pill	Expectations, State anxiety, Trait anxiety
Stegen et al. (2000) a,b,d,e	P	Healthy psychology students (44, U/K, 27.3)	Breathing trial with room air	Negative affect, Social desirability
Strohle (2000) e	P	Healthy adults and patients with panic disorder (U/K, 33.5, 56.6)	Sham panic disorder trigger	Gender, Clinical condition
Sullivan et al. (2008) c,e	P	Patients with neuropathic pain (24, 54.7, 62.5)	Sham cream treatment	Pain catastrophising
Vase et al. (2013) e	P	Patient with pain due to tooth removal (U/K, 25.5, 47.5)	Sham acupuncture	Expectations
Wendt et al. (2014) e	P	Healthy males (24, 25.0, 100.0)	Sham immuno-suppressive capsule	Genes

Note: RCT = Randomised Controlled Trial, Non RCT = Non randomised controlled trial, P = Prospective design, B = Between subjects design, W = Within subjects design, U/K= Unknown, *italicised* = not directly given but has been extrapolated from the available data, F = Female, M = Male, ns = non-significant, UPDRS = Unified Parkinson's disease rating scale, RF-EMF = Radio frequency electromagnetic fields, EMF = Electromagnetic fields, **a** = high risk for selection bias, **b** = high risk for confounding factors, **c** = high risk for insufficient follow-up, **d** = high

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

risk for low generalisability, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor , N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 3. The effect of learning on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Andre-Obadia et al. (2011) b,e	RCT (W)	Chronic neuropathic pain patients (45, 55.0, 37.8)	Sham rTMS	1. Prior experience: a. Sham rTMS before active rTMS (20); b. Sham rTMS after successful active rTMS (12); c. Sham rTMS after ineffective active rTMS (13)	i. Mean pain intensity ($c > a > b$) "placebo sessions tended to worsen pain when following an unsuccessful rTMS" – no statistics given ii. Combined pain assessment "Comparable results obtained" – no statistics given	No significant interactions with baseline pain ratings
Bayer et al. (1998) a,e	RCT (B)	Job seekers (62, U/K, 82.0)	Sham electrical shock	1. Prior experience: a. Exposed to two physical pain induction procedures prior to sham stimulation (32); b. Warned of pain and received sham stimulation. They were not exposed to any prior pain induction (30)	i. Subjects reporting pain (ns) ii. Mean maximal pain rating (ns) iii. Subjects reporting pain over analgesic threshold (ns) iv. Pain intensity over time (ns)	Prior experience x Dose i-iii. Not assessed iv. Increased with increasing stimulator settings for those in condition a ($p < .01$), ns for those in condition b
Dalton (1999) e	RCT (B)	Healthy volunteers (180, 31.7, 49.4)	Odours	1. Odours: a. Pleasant smelling methyl salicylate (60); b. neutral smelling isobornyl acetate (60); c. Foul smelling butanol (60)	i. Symptom reports (ns)	Odours x Verbal symptom suggestions i. "Highest for those exposed to butanol following negative suggestions, lowest for those exposed to methyl salicylate following positive suggestions" No other interactions assessed
De Peuter et al. (2005) e	RCT (W)	Asthma patients and healthy controls (40, 23.9, 52.5)	Sham inhaler	1. Conditioning: a. One sham inhaler previously paired with CO ₂ challenge; b. One sham inhaler previously paired with O ₂	i. Total symptom score ($a > b$, $p < .01$) ii. Obstruction (ns) iii. Dyspnea ($a > b$, $p < .01$) iv. Fatigue (ns) v. Hyperventilation (ns) vi. Anxiety (ns) vii. Irritation (ns)	No significant interaction with clinical condition No other interactions assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Devriese et al. (2000) a,e	Non RCT (B+W)	Healthy students (56, U/K, 41.1)	Odours	<p>1. Odour: Within each group participants were exposed to foul smelling ammonia and pleasant smelling niaouli</p> <p>2. Conditioning: a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (28); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (28)</p> <p>3. Generalisation: Within each group participants were exposed to a. a new foul smelling odour butyric acid; b. a new foul smelling odour acetic acid; and c. a new pleasant smelling odour citric aroma</p>	<p>i. Total symptom score (1 ns; 2 ns; 3 a>b+c, p<.001, b vs c ns)</p> <p>ii. Arousal (1 ns; 2 ns; 3 ns)</p> <p>iii. Respiratory (1 ns; 2 ns; 3 ns)</p> <p>iv. Cardiac (1 ns; 2 ns; 3ns)</p> <p>v. Tingling (1 ns; 2 ns; 3 ns)</p> <p>vi. Unclassified (1 ns; 2 ns; 3 ns)</p> <p>vii. Dummy (1 ns; 2 ns; 3 ns)</p>	<p>Conditioning x Odour</p> <p>i+iii. Higher in response to odour paired with CO2 but only when the odour was ammonia (p<.05)</p> <p>ii, iv-vii. ns</p> <p>Conditioning x Odour x Negative affect</p> <p>i+iii. Higher in response to odour paired with CO2 but only when the odour was ammonia and participants had high negative affect (p<.05)</p> <p>ii, iv-vii. ns</p> <p>Odour x Negative affect x Generalisation</p> <p>i+iii. Higher in response to butyric and acetic acid than citric aroma when ammonia was paired with CO2 and participants had high negative affect (p<.05)</p> <p>iv. Higher in response to butyric acid than acetic acid or citric aroma when ammonia was paired with CO2 and participants had high negative affect (p<.05)</p> <p>ii+ v-vii. Ns</p> <p>No other interactions assessed</p>
Devriese et al. (2004) a,e	Non RCT (B+W)	Healthy students (53, U/K, U/K)	Odours	<p>1. Odour: Within each group participants were exposed to foul smelling ammonia and foul smelling butyric acid</p> <p>2. Conditioning: a. Ammonia paired with CO2 breathing task, butyric acid paired with room air breathing task (28); b. Ammonia paired with room air breathing task, butyric acid paired with CO2 breathing task (25)</p>	<p>i. Total symptom score (1 ns; 2 ns)</p> <p>ii. Arousal (1 ns; 2 ns)</p> <p>iii. Respiratory (1 ns; 2 ns)</p> <p>iv. Cardiac (1 ns; 2 ns)</p> <p>v. Tingling (1 ns; 2 ns)</p> <p>vi. Unclassified (1 ns; 2 ns)</p> <p>vii. Dummy (1 ns; 2 ns)</p>	<p>Odour x Conditioning</p> <p>i. Higher in response to butyric acid than ammonia when butyric acid paired with room air (p<.01)</p> <p>ii-vii. ns</p> <p>Perceived cue odour x Odour</p> <p>i. Higher to butyric acid than ammonia when butyric acid was thought to have been paired with CO2 (p<.05)</p> <p>ii-vii. ns</p> <p>No other interactions assessed</p>

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Devriese et al. (2006)	RCT (B+W)	Psychology students (40, U/K, 0.0)	Odour	1. Odour: Within each group participants were exposed to foul smelling ammonia and foul smelling acetic acid 2. Conditioning: a. Ammonia paired with CO2 breathing task, acetic acid paired with room air breathing task (20); b. Ammonia paired with room air breathing task, acetic acid paired with CO2 breathing task (20)	i. Total symptom score (1 ns; 2 higher for the odour paired with CO2 than the odour paired with room air, $p<.05$) ii. Arousal (1 ns; 2 ns) iii. Respiratory (1 ns; 2 ns) iv. Cardiac (1 ns; 2 ns) v. Tingling (1 ns; 2 ns) vi. Unclassified (1 ns; 2 ns) vii. Dummy (1 ns; 2 ns)	Not assessed
Dinnerstein and Halm (1970) c,e	RCT (B)	Male students (80, U/K, 100.0)	Sham arousal liquid	1. Prior experience: a. Received aspirin prior to sham (40); b. Received lactose prior to sham (40)	i. Friendly (ns) ii. Aggressive (ns) iii. Sleepy (ns) iv. Dizzy (ns) v. Unhappy (ns) vi. Clear thinking (ns)	Prior experience x Arousal suggestion i. Lower in condition a than condition b under energiser suggestion, reverse under tranquiliser suggestion ($p<.05$) ii-iv. ns v+vi. Higher in condition a than condition b under energiser suggestion, reverse under tranquiliser suggestion ($ps<.05$)
Flaten and Blumenthal (1999) e	RCT (W)	Healthy coffee drinkers (21, 24.8, 61.9)	Decaffeinated solution	1. Association: a. Given in an orange juice solution; b. Given in a coffee solution	i. Stress score (ns) ii. Arousal score ($b>a$, $p<.05$) iii. Alertness score ($b>a$, $p<.05$)	N/A
Lorber et al. (2007) e	RCT (B)	Students without upper respiratory conditions (86, U/K, 40.7)	Sham environmental toxin	1. Social observation: a. Told inhaled substance has been reported to produce symptoms and observed a female confederate inhale and display symptoms (U/K) b. As above but no observation of confederate (U/K) c. Did not inhale the substance and observed a female confederate inhale and display	i. Specified verbal symptom ratings: headache, nausea, itchy skin, drowsiness (ns) ii. Other verbal symptom ratings: watery eyes, scratchy throat, chest tightness, and breathing difficulty (ns)	Observation x Gender i. Higher in observation compared to no observation conditions when participants are female ($p<.05$) ii. ns

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				symptoms (U/K) d. As above but no observation of confederate (U/K)		
Mazzoni et al. (2010) e	RCT (B)	Healthy students (120, 20.7, 50.0)	Sham environme ntal toxin	1. Social observation: a. Observed a male/female confederate inhale the substance and display symptoms (60) b. Did not observe a male/female confederate inhale the substance and display symptoms (60)	i. Specified verbal symptom ratings: headache, nausea, itchy skin, drowsiness (a>b, p<.001) ii. Other verbal symptom ratings: watery eyes, scratchy throat, chest tightness, and breathing difficulty (ns)	Not assessed
Meulders et al. (2010) a,e	Non RCT (B+W)	Healthy adults (58, 22.0, 48.3)	Odours	1. Odour: Within each group participants were exposed to foul smelling ammonia and foul smelling butyric acid 2. Conditioning: a. Ammonia paired with CO2 breathing task, butyric acid paired with room air breathing task (29); b. Ammonia paired with room air breathing task, butyric acid paired with CO2 breathing task (29)	i. Total symptom score (1 ns; 2 ns) ii. Arousal (1 ns; 2 ns) iii. Respiratory (1 ns; 2 ns) iv. Cardiac (1 ns; 2 ns) v. Tingling (1 ns; 2 ns) vi. Unclassified (1 ns; 2 ns) vii. Dummy (1 ns; 2 ns)	Conditioning x Ability to predict i. Higher in response to odours which had been paired with CO2 compared to room air when participants were able to predict which odour had caused the most symptoms (p<.05) ii-vii. ns Conditioning x Ability to predict x Odour i-iv+vii. Higher in response to butyric acid which had been paired with CO2 compared to room air when participants were able to predict which odour had caused the most symptoms (i-iv, p<.01; vii, p<.05) v+vi. ns No other interactions assessed
Mikalsen et al. (2001) a,b,e	RCT (W)	Student coffee drinkers (21, 25.9, 66.7)	Sham coffee	1. Association: a. Given in a juice solution (U/K); b. Given in a coffee solution (U/K)	i. Alertness VAS score (b>a, p<.05) ii. Contentedness VAS score (b>a, p=.02) iii. Calmness VAS score (ns)	Not assessed
Papoiu et al. (2011) e	RCT (W)	Healthy volunteers and patients with	Sham histamine	1. Social Observation: a. Watched a 5 minute video of people scratching their left	i. Average itch intensity rating (a>b for patients p=.027; ns for healthy volunteers)	No significant interactions with gender

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

		atopic dermatitis (25, U/K, 44.0)		forearm b. Watched a 5 minute video of the same persons in the scratching video but sitting idle.	ii. Scratching behaviour (a>b for patients p=.002; ns for healthy volunteers)	
Stegen et al. (1998) e	RCT (W)	Healthy psychology students (72, U/K, 48.6)	Breathing trial with room air	1. Prior experience: a. Room air breathing trial before 7.5% CO2 challenge; b. Room air breathing trial after 7.5% CO2 challenge	i. Total symptom score (ns) ii. General arousal (a>b, p<.001) iii. Respiration (a>b, p<.001) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Gastrointestinal (ns) viii. Dizziness (ns)	Prior experience x Negative affect i+ii. Participants scoring high on negative affect reported more complaints than participants with low negative affect in condition a (i, p<.001; ii, p<.005) iii-viii. ns
Van den Bergh et al. (1995) a,e	Non RCT (B+W)	Healthy students (28, U/K, 50.0)	Odours	1. Odour: Within each group participants were exposed to foul smelling ammonia and pleasant smelling niaouli 2. Conditioning: a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (14); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (14)	i. Total symptom score (1 ns; 2 ns) ii. Arousal (1 ns; 2 ns) iii. Respiratory (1 ns; 2 ns) iv. Cardiac (1 ns; 2 ns) v. Tingling (1 ns; 2 ns) vi. Unclassified (1 ns; 2 ns)	Conditioning x Odour i-iii+vi. Higher in response to odour paired with CO2 but only when that odour was ammonia (i, p<.05; ii, p<.001; iii, p<.02; vi, p<.05) iv+v. ns No other interactions assessed
Van den Bergh et al. (1997) a,e	Non RCT (B+W)	Psychosomatic patients (28, 36.0, 50.0)	Odours	1. Odour: Within each group participants were exposed to foul smelling ammonia and pleasant smelling niaouli 2. Conditioning: a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (14); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (14)	i. Total symptom score (1 ns; 2 higher for odours that had been paired with CO2 compared to room air; 3 ns) ii. Arousal (1 ns; 2 ns; 3 ns) iii. Respiratory (1 ns; 2 ns; 3 ns) iv. Cardiac (1 ns; 2 ns; 3 ns) v. Tingling (1 ns; 2 ns; 3 ns) vi. Unclassified (1 ns; 2 ns; 3 ns) vii. Dummy (1 ns; 2 ns; 3 ns)	Conditioning x Odour i+iii+vi. Higher for odour paired with CO2 but only when that odour was ammonia (i+iii, p<.001; vi, p<.005) ii+iv+v+vii. Ns No other interactions assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				3. Generalisation: Within each group participants were exposed to a new foul smelling odour Ichtyol and new pleasant smelling odour Rose		
Van den Bergh et al. (1998) e	RCT (B+W)	Healthy adults (56, 42.5, 50.0)	Odours	<p>1. Odour: Within each group participants were exposed to foul smelling ammonia and pleasant smelling niaouli</p> <p>2. Conditioning: a. Ammonia paired with CO2 breathing task, Niaouli paired with room air breathing task (28); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (28)</p> <p>3. Generalisation: Within each group participants were exposed to a new foul smelling odour Ichtyol and new pleasant smelling odour Rose</p>	<p>i. Total symptom score (1 ns; 2 ns; 3 higher in response to ichtyol than rose odour, $p<.005$)</p> <p>ii. Arousal (1 ns; 2 ns; 3 ns)</p> <p>iii. Respiratory (1 ns; 2 ns; 3 ns)</p> <p>iv. Cardiac (1 ns; 2 ns; 3 ns)</p> <p>v. Tingling (1 ns; 2 ns; 3 ns)</p> <p>vi. Unclassified (1 ns; 2 ns; 3 ns)</p> <p>vii. Dummy (1 ns; 2 ns; 3 ns)</p>	Conditioning x Odour x Self-awareness i+ii+iv+vi+vii. Higher in response to the odour paired with CO2 when this was ammonia and participants had not been distracted (i+ii, $p<.001$, iv+vi+vii, $p<.002$) ii+v. ns No other interactions assessed
Van den Bergh et al. (1999) a,e	Non RCT (B+W)	Healthy students (64, U/K, 25.0)	Odours	<p>1. Odour: Within each group participants were exposed to foul smelling ammonia and foul smelling butyric acid</p> <p>2. Conditioning: a. Ammonia paired with CO2 breathing task, butyric acid paired with room air breathing task (32); b. Ammonia paired with room air breathing task, butyric acid paired with CO2 breathing task (32)</p>	<p>i. Total symptom score (1 ns; 2 higher in response to odour paired with CO2 compared to room air, $p<.001$)</p> <p>ii. Arousal (1 ns; 2 ns)</p> <p>iii. Respiratory (1 ns; 2 higher in response to odour paired with CO2 compared to room air, $p<.001$)</p> <p>iv. Cardiac (1 ns; 2 ns)</p> <p>v. Tingling (1 ns; 2 ns)</p> <p>vi. Unclassified (1 ns; 2 higher in response to odour paired with CO2 compared to room air, $p<.001$)</p>	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Van Diest et al. (2006) e	RCT (B+W)	Students (28, U/K, 21.4)	Odours	1. Odour: Within each group participants were exposed to foul smelling ammonia and foul smelling acetic acid 2. Conditioning: a. Ammonia paired with hypocapnic over breathing trial, acetic acid paired with normocapnic over breathing trial (13); b. Ammonia paired with normocapnic over breathing trial, acetic acid paired with hypocapnic over breathing trial (15)	vii. Dummy (1 ns; 2 ns) i. Paresthesia (1 ns; 2 ns) ii. Cerebral (1 ns; 2 ns) iii. Cardiac (1 ns; 2 ns) iv. Gastrointestinal (1 ns; 2 ns) v. Respiratory (1 ns; 2 ns) vi. Anxiety (1 ns; 2 ns) vii. Neuropsychological (1 ns; 2 ns) viii. Unclassified (1 ns; 2 ns) ix. Dummy (1 ns; 2 ns) x. Feeling unreal (1 ns; 2 higher for odour paired with hypocapnic compared to normocapnic over breathing)	Not assessed
Winters et al. (2001) Exp 1 a,e	Non RCT (B)	Psychology students (50, U/K,U/K)	Ammonia	1. Conditioning: a. Odour + CO2 trials and room air trials (10); b. Odour trials and CO2 trials (10); c. Odour trials, CO2 trials, odour + CO2 trials, room air trials (10); d. odour trials, room air trials (10); e. CO2 trials, room air trials (10)	i. Symptom reports (ns)	N/A
Winters et al. (2001) Exp 2 a,e	Non RCT (B)	18-30 year olds (40, U/K,U/K)	Odour	1. Odour: a. Foul smelling ammonia (20); b. Pleasant smelling niaouli (20) 2. Conditioning: a. Odour + CO2 trials and room air trials (20); b. Odour trials and CO2 trials (20)	i. Symptom reports (1 ns; 2 ns)	Not assessed
Winters et al. (2003) a,e	Non RCT (B+W)	18-30 year olds (32, U/K,15.6)	Odour	1. Odour: Within each group participants were exposed to foul smelling ammonia and pleasant smelling niaouli 2. Conditioning: a. Ammonia paired with CO2 breathing	i. Total symptom score (1 ns; 2 higher in response to odour paired with CO2 compared to room air, p<.05) ii. Arousal (1 ns; 2 ns) iii. Respiratory (1 ns; 2 ns)	Conditioning x Odour i. Higher in response to the odour paired with CO2 when this was ammonia (p<.05) ii-vii. ns Conditioning x Verbal symptom

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

task, Niaouli paired with room air breathing task (16); b. Ammonia paired with room air breathing task, Niaouli paired with CO2 breathing task (16)	iv. Cardiac (1 ns; 2 ns) v. Tingling (1 ns; 2 ns) vi. Unclassified (1 ns; 2 ns) vii. Dummy (1 ns; 2 ns)	suggestions i+iii+iv+vi. Higher following odour which was paired with CO2 when given symptom suggestions (i, $p<.01$; iii+iv+vi, $p<.05$) ii+v+vii. ns No other interactions assessed
---	--	--

Note: RCT = Randomised Controlled Trial, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, italicised = not directly given but has been extrapolated from the available data, rTMS = Repetitive Transcranial Magnetic Stimulation, CO2 = Carbon dioxide, O2 = Oxygen, **a** = high risk random sequence generation bias, **b** = high risk allocation concealment bias, **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 4. The effect of perceived dose manipulation on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Bayer et al. (1991) e	RCT (B+W)	Unemployed Men (100, U/K, 100.0)	Sham electrical shock	a. Within each group the stimulator setting increased from 0 to 80 mA	i. Mean pain ratings (increased with greater sham stimulation, $p<.001$)	No significant interactions with symptom suggestion
Bayer et al. (1998) a,e	RCT (B+W)	Job seekers (62, U/K, 82.0)	Sham electrical shock	a. Within each group the stimulator setting increased in steps of 10 every 5 minutes till it reached 50	i. Number of subjects reporting pain (ns) ii. Mean maximal pain rating (ns) iii. Subjects reporting pain over analgesic threshold (ns) iv. Pain intensity rating over time (Increased, difference between stimulator settings, $p<.01$)	Perceived dose x Prior experience i-iii. Not assessed iv. Increased with increasing stimulator settings for those who experienced prior pain ($p<.01$), ns for those with no prior pain
Flaten et al. (2003) a,b,e	W	Coffee drinkers (20, U/K, 50.0)	Sham coffee	a. Participants were first given one cup and then a second	i. Alertness, contentedness, calmness, arousal, and stress VAS scores (ns)	Not assessed
Jensen and Karoly (1991) e	RCT (B+W)	Students (86, U/K, 45.3)	Sham sedative pill	a. Suggestions of a high dose or low dose were counterbalanced across each group	i. General placebo response rating (ns) ii. Adjective symptom checklist score (ns)	No significant interactions with social desirability or gender
Kirsch and Weixel (1988) e	RCT (B)	Student coffee drinkers (U/K, 19.3, 31.0)	Sham coffee	a. 1 tsp (U/K) b. 2 tsps (U/K) c. 3 tsps (U/K) d. 5 tsps (U/K) e. 8 tsps (U/K)	i. Mean change in alertness (ns) ii. Mean change in relaxation (ns) iii. Mean change in tension ("significant linear increase in tension as a function of dose", $p<.03$)	Perceived dose x Likelihood suggestion i+iii. Increased with increasing dose in the deceptive group, decreased in the double-blind group (i, $p<.02$; iii, $p<.04$) ii. ns
Szemerszky et al. (2010) a,b,c,e	W	Healthy students (40, 22.8, 27.5)	Sham EMF	a. Told it would be weak b. Told it would be strong	i. Overall symptom score ($b>a$, $p<.001$)	Not assessed

Note: RCT = Randomised Controlled Trial, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, EMF = Electromagnetic Field, tsp = teaspoon, italicised = not directly given but has been extrapolated from the available data, **a** = high risk random sequence generation bias, **b** = high risk allocation concealment bias, **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 5. The effect of self-awareness manipulation on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Geers, Helfer, et al. (2005) e	RCT (B)	Healthy students (54, 21.0, 29.6)	Sham over-the-counter pill	a. Told to attend to any symptoms experienced (27) b. Not given any such instructions (27)	i. Affect questionnaire - Anxiety, restlessness, relaxed, irritable, and perspiration (ns)	No significant interactions with age, gender, likelihood suggestion or optimism
Geers et al. (2006) e	RCT (B)	Healthy students (54, U/K, 31.5)	Sham over-the-counter pill	a. Told to closely monitor feelings/bodily sensations (27) b. Not given any such instructions (27)	i. Overall placebo symptom index: anxiety, nausea, pleasant feelings, perspiration, and perceived pill effect (ns)	Self-awareness x Likelihood suggestion i. Higher in condition a participants with deceptive suggestions than condition a participants with control suggestions ($p<.01$) and condition a participants with double-blind suggestions ($p=.02$). In addition it is higher in condition a participants with deceptive suggestions than condition b participants with deceptive suggestions ($p=.02$)
Gibbons et al. (1979) a,e	RCT (B)	Female students (38, U/K, 0.0)	Sham drug	a. Mirror was facing participants (19) b. Mirror was not facing participants (19)	i. Perceived arousal ($b>a$, $p<.01$) ii. Salient symptom checklist (ns) iii. Non salient symptom checklist (ns)	Self-awareness x Symptom suggestion i +iii. ns ii. Lower in condition a than condition b when participants were misinformed ($p<.03$)
Van den Bergh et al. (1998) e	RCT (B+W)	Healthy adults (56, 42.5, 50.0)	Odours	a. Told to count lower tones and disregard higher tones (28) b. Told to ignore tones (28)	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	Self-awareness x Conditioning i+ii+iv+vi+vii. Higher in response to the odour paired with CO ₂ when this was ammonia and participants were in condition b (i+ii, $p<.001$, iv+vi+vii, $p<.002$) ii+v. ns No other interactions assessed

Note: RCT = Randomised Controlled Trial, B = Between subjects design, U/K= Unknown, ns = non-significant, CO₂ = Carbon dioxide, **a** = high risk random sequence generation bias, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 6. The effect of type of administration manipulation on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Goldman et al. (1965) a,b,e	Non RCT (B)	Male veterans with schizophrenia (64, 44.0, 100.0)	Sham arousal treatment	a. Received sugar pill (32) b. Received saline injection (32)	i. Reported symptoms identified through interviews (ns) ii. Reported drug effect identified through interviews (ns) iii. Ward activity (ns)	Not assessed
Kaptchuk et al. (2006)	RCT (B)	Adults with distal pain in the arms (266, 36.7, 45.9)	Sham treatment	a. Received sham acupuncture (133) b. Received placebo pill (133)	i. Frequency of one or more side effects reported (ns)	N/A

Note: RCT = Randomised Controlled Trial, Non RCT = Non-randomised controlled trial B = Between subjects design, ns = non-significant, **a** = high risk random sequence generation bias, **b** = high risk allocation concealment bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 7. The effect of verbal suggestions on performance on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Harrell and Juliano (2009) c	RCT (B)	Adult non-smoking coffee consumers (30, 22.6, 22.0)	Sham coffee	a. Told caffeine enhances performance (15) b. Told caffeine impairs performance (15)	i. Sum of ten self-reported symptoms (ns) ii. Profile of mood states score (ns)	N/A
Harrell and Juliano (2012) c,e	RCT (B)	Adult smokers (43, 28.7, 67.4)	Sham cigarette	a. Told cigarette enhances performance (20) b. Told cigarette impairs performance (23)	i. Cigarette evaluation scale - cravings score (b>a, p=.02)	No significant interaction with gender
Nevelsteen et al. (2007) e	RCT (B)	Healthy males (59, 48.4, 100.0)	Sham magnetic field	a. Told magnetic fields enhance cognitive performance (15) b. Told magnetic fields impair cognitive performance (15) c. Told magnetic fields have no effect on cognitive performance (14) d. Not exposed to sham magnetic field and received no information (15)	i. Subjective vigilance feelings (ns) ii. Profile of mood states score (ns) iii. 24 Physical symptoms scale (ns)	Not assessed

Note: RCT = Randomised Controlled Trial, B = Between subjects design, ns = non-significant, italicised = not directly given but has been extrapolated from the available data, **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 8. The effect of verbal suggestions of likelihood on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Geers et al. (2006) e	RCT (B)	Healthy students (54, U/K, 31.5)	Sham over-the-counter pill	a. Told the pill had unpleasant side effects (18) b. Told they may or may not receive the active drug (19) c. Told they would ingest an inactive drug (17)	i. Overall placebo symptom index: anxiety, nausea, pleasant feelings, perspiration, and perceived pill effect (ns)	Likelihood suggestion x Self-awareness i. Higher in condition a than condition c ($p<.01$) or condition b participants ($p=.02$) when participants told to attend to sensations. Higher in participants told to attend to sensations than those not given such instructions when in condition a ($p=.02$)
Geers et al. (2011) e	RCT (B)	Healthy students (102, 20.5, 21.6)	Sham caffeine capsule	a. Told it contained 250mg of caffeine (34) b. Told they may or may not be ingesting 250mg of caffeine (34) c. Not given the capsule and received no caffeine expectation (34)	i. Placebo response index-anxious, sluggish, energized, calm, irritated, lazy, relaxed, and excited ($a>b$, $p<.05$; $a>c$, $p<.01$; b vs c , ns)	No significant interactions with gender
Geers, Helfer, et al. (2005) e	RCT (B)	Healthy students (54, 21.0, 29.6)	Sham over-the-counter pill	a. Told the pill had unpleasant side effects (18) b. Told the pill would make them feel either unpleasant or was an inactive substance (18) c. Told they would ingest an inactive pill (18)	i. Affect questionnaire - Anxiety, restlessness, relaxed, irritable, and perspiration (ns)	Likelihood suggestion x Optimism i. Higher in pessimists than optimists when in condition a than condition c ($p<.05$), condition a vs b, b vs c (ns) No significant interaction with age, gender or self-awareness
Kirsch and Weixel (1988) e	RCT (B)	Student coffee drinkers (100, 19.3, 31.0)	Sham coffee	a. Told they would receive coffee (U/K) b. Told they may or may not receive caffeinated coffee (U/K) c. No beverage, waited for 20 minutes (U/K)	i. Mean change in alertness ($a+b>c$, $p<.003$; $a>b$, U/K) ii. Mean change in relaxation (ns) iii. Mean change in tension (ns)	Likelihood suggestion x Perceived dose i+iii. Increased with increasing dose in condition a, decreased in condition b (i, $p<.02$; iii, $p<.04$) ii. ns
Ossege et al. (2005)	RCT (B)	Healthy volunteers (60, 27.6, 40.0)	Placebo drug	a. Told it was an active medication (30) b. Told there was a 50% chance that it was a placebo or active medication (30)	i. Number of adverse events (ns) ii. Severity of adverse events (ns)	N/A

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Tippens et al. (2014) e	RCT (B)	Obese adults (79, 49.4, 10.4)	Sham weight loss supplement	a. Told they would be given an active weight loss supplement (27) b. Told they would be randomly assigned to either the active or placebo supplement (28) c. Only received lifestyle education (24)	i. Number of recorded adverse events (a+b>c, p <.001; a vs b, ns)	N/A
Walach and Schneider (2009) Exp 1	RCT (B)	Healthy adult coffee drinkers (60, 32.3, 23.3)	Sham coffee	a. Told it was caffeine (15) b. Told it could be placebo or caffeine (15) c. Told it could be placebo or caffeine (15) d. Received no beverage (15)	i. Mean change in mood (ns) ii. Mean change in calmness (ns) iii. Mean change in alertness (ns)	Not assessed
Walach et al. (2001)	RCT (B)	Coffee drinkers (157, 28.1, 34.0)	Sham coffee	a. Told they would receive a placebo (41) b. Told they would receive coffee (39) c. Told they may receive real coffee or decaffeinated coffee (39) d. No substance or instruction given (38)	i. General wellbeing score (a>d, p<.0004; all other comparisons, ns)	No significant interaction with experimenter expectations
Walach et al. (2002)	RCT (B)	Coffee drinkers (159, 25.5, 58.0)	Sham coffee	a. Told they would receive a placebo (39) b. Told they would receive coffee (40) c. Told they may receive real coffee or decaffeinated coffee (40) d. No substance or instruction given (40)	i. General wellbeing score (ns)	No significant interaction with Symptom suggestion

Note: RCT = Randomised Controlled Trial, B = Between subjects design, U/K= Unknown, ns = non-significant, italicised = not directly given but has been extrapolated from the available data, **e** = did not mention an a priori sample size calculation, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 9. The effect of verbal suggestions of arousal on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean Age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Angelucci and Pena (1997) e	RCT (B)	Student caffeine consumers (148, U/K, 23.0)	Sham coffee	a. Given coffee with no expectations (37) b. Given coffee with low arousal expectations (37) c. Given coffee with high arousal expectations (37) d. no coffee and no expectations (37)	I. Stimulation/alertness ($b < a + c + d$, $p < .001$; all other comparisons, ns) ii. Anxiety/irritability (ns) iii. Subjective symptoms (ns)	Not assessed
Brodeur (1965) e	RCT (B)	Healthy senior students (45, U/K, 91.1)	Sham arousal capsule	a. Told it was a stimulant (15) b. Told it was a tranquilizer (15) c. No suggestion (15)	i. Arousal score (ns)	N/A
Dinnerstein and Halm (1970) e	RCT (B)	Male students (80, U/K, 100.0)	Sham arousal liquid	a. Told it was an energizer (40) b. Told it was a tranquilizer (40)	i. Friendly ($b < a$, $p < .025$) ii. Aggressive ($b < a$, $p < .025$) iii. Sleepy ($b > a$, $p < .001$) iv. Dizzy (ns) v. Unhappy (ns) vi. Clear thinking (ns)	Arousal suggestion x Prior experience i. Lower after aspirin than lactose under condition a, reverse under condition b ($p < .05$) ii-iv. ns v+vi. Higher after aspirin than lactose under condition a, reverse under condition b ($p < .05$)
Flaten (1998) e	RCT (B)	Healthy students (48, U/K, 35.4)	Sham arousal drink	a. Told you will feel relaxed and sleepy (16) b. Told you will feel alert and a little stress (16) c. Told you will take an inactive drug (16)	i. Subjective stress score ($a < b$, $p < .05$; $a < c$, $p < .05$; b vs c , ns) ii. Subjective arousal score ($a < b$, $p < .05$; all other comparisons, ns)	N/A
Flaten et al. (1999) e	RCT (B)	Healthy volunteers in non-health professions (34, U/K, 54.5)	Sham arousal capsule	a. The drug will make you feel relaxed (11) b. The drug will make you feel alert (12) c. You will receive capsules that contain a prescription drug (11)	i. Sleep-wake dimension score (ns) ii. Relaxed-tense dimension score (tense score $b > a$, $p = .041$; all other comparisons, ns)	N/A

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Geers, Weiland, et al. (2005) e	RCT (B)	Healthy student (57, U/K, 35.1)	Sham caffeine pill	a. Told they were given caffeine (U/K) b. No mention of caffeine (U/K)	i. Number of caffeine symptoms reported (a>b, p=.03)	Arousal suggestion x Cooperation prime i. Higher in condition a than b when primed for cooperation (p=.02), when not primed for cooperation there is no significant difference between condition a+b No significant interaction with caffeine consumption
Goldman et al. (1965) e	Non RCT (B)	Male veterans with Schizophrenia (64, 44.0, 100.0)	Sham arousal treatment	a. Told it would heighten their ward activity (32) b. Told it would lower their ward activity (32)	i. Reported symptoms identified through interviews (U/K) ii. Reported drug effect identified through interviews (b>a, p<.001) ii. Ward activity (ns)	Not assessed
Higuchi et al. (2002) e	RCT (B)	Healthy volunteers (30, 21.2, 40.0)	Fragrance (Lavender or Jasmine)	a. Told it was relaxing (10) b. Told it was stimulating (10) c. No information given (10)	i. Relaxed scores (ns) ii. Stimulant scores (Lavender, ns; Jasmine, c>a, p<.05; other comparisons U/K) iii. Stress reduced scores (Lavender, a>c, p<.05, a vs b, ns, b vs c, U/K; Jasmine, ns)	N/A
Kuenzel et al. (2012) e	RCT (B)	English speaking students (148, 21.7, 18.2)	Herbal infusion tea	a. Told it would make them feel relaxed (45) b. Told it would make them feel active (53) c. No information given (50)	i. Symptom ratings (ns)	N/A
Lotshaw et al. (1996) e	RCT (B)	Male student coffee drinkers (50, U/K, 100.0)	Sham coffee	a. Told coffee received decaffeinated (25) b. Told decaffeinated received decaffeinated (25)	i. Profile of mood states score (ns) ii. Effects of coffee received score (a>b, "subjects who were told they were receiving caffeine rated the effects to be significantly greater than subjects who were told they were receiving decaffeinated")	N/A

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Mikalsen et al. (2001) e	RCT (W)	Student coffee drinkers (21, 25.9, 66.7)	Sham coffee	a. Told it was caffeine b. Told it was not caffeine	i. Alertness VAS score (ns) ii. Contentedness VAS score (ns) iii. Calmness VAS score (a<b, p<.05)	Not assessed
Mrna and Skrivanek (1985) a,b,e	W	Healthy volunteers (21, 17.0, 47.6)	Sham arousal drug	a. Told it was a new doping drug undetectable by anti-doping tests b. Told it was to relax pre-restart states	i. Observed behaviour ("after the doping drug students were lively and talkative whereas after the sedative they sat quietly and some fell asleep")	Not assessed
Penick and Fisher (1965) a,b,c,e	W	Healthy medical students (14, U/K, U/K)	Sham arousal drug	a. Told they would receive a stimulant drug b. Told they would receive a sedative drug	i. Overall arousal score (ns)	N/A
Schneider et al. (2006) c,e	RCT (B)	Healthy Adults (45, 31.0, 22.2)	Sham coffee	a. Told they were to consume decaffeinated coffee (15) b. Told they were to consume regular coffee (15). c. Informed they would receive no beverage and no instructions (15)	i. Mood (ns) ii. Alertness (b>c, p=.04; all other comparisons, ns) iii. Calmness (ns)	N/A
Slanska et al. (1974) a,e	Non RCT (B)	Medical students (33, U/K, U/K)	Salt solution	a. Told it was a stimulant (17) b. Told it was a sedative (16)	i. Perceived effect ("stimulation suggestion was effective in 12% of placebo reactors, sedation suggestion was effective in 25%") ii. Vigilance ("Sedative suggestion decreased alertness") iii. Fatigue ("b>a, statistically significant increase in fatigue") iv. Tension v. Relaxation	Not assessed
Walach and Schneider (2009)- Exp 2	RCT (B)	Healthy adults coffee drinkers (30, 29.9, 33.3)	Sham coffee	a. Told it was caffeine (15) b. Received no beverage (15)	i. Mean change in mood (ns) ii. Mean change in calmness (ns) iii. Mean change in alertness (a>b, d=0.64)	Not assessed

Note: RCT = Randomised Controlled Trial, Non RCT = Non randomised controlled trial, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, italicised = not directly given but has been extrapolated from the available data, **a** = high risk random sequence generation bias, **b** = high risk allocation concealment bias **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 10. The effect of verbal suggestions of symptoms on symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Bayer et al. (1991) e	RCT (B)	Unemployed Men (100, U/K, 100.0)	Sham electrical shock	a. Told they would receive a safe but often painful undetectable current (60) b. Were assured there would be no shocks (40)	i. Mean pain rating (a>b, p<.01) ii. Frequency of pain reports (a>b, p<.05)	No significant interactions with dose
Benedetti et al. (1997) e	RCT (B)	Video assisted thoracoscopy patients (36, 53.7, 66.1)	Sham treatment	a. Open injection that it would increase pain (18) b. Hidden injection (18)	i. Pain intensity rating (a>b, p<.005)	N/A
Colagiuri et al. (2012) e	RCT (B)	Students experiencing sleep difficulty (82, 20.2, 22.0)	Sham sleeping pill	a. Treatment might cause one side effect (29) b. Treatment might cause four side effects (23) c. No warning about side effects (30)	i. Free side effect report (ns) ii. Change in appetite report (ns) iii. Severity of change in appetite (ns) iv. Cued report side effect (c>a+b, restlessness p=.04, poor concentration, p=.001; a vs b, ns) v. Severity of cued side effects ("trend" for c>a+b; a vs b, ns)	N/A
Crichton et al. (2014) e	RCT (B)	Students (54, U/K, 37.0)	Sham infrasound	a. TV footage detailing symptomatic experiences attributed to wind farms (27) b. TV footage with experts stating wind farms would not cause symptoms (27)	i. Total symptom change score (a>b, p<.01) ii. Total symptom change severity score (a>b, p<.001)	N/A
Dalton (1999) e	RCT (B)	Healthy volunteers (180, 31.7, 49.4)	Odours	a. Told they would have relaxing effects (60) b. Told they were industrial solvents (60) c. Told they were approved for olfactory research (60)	i. Symptom reports (b>a+c, p<.05)	Symptom suggestions x Odour i. "Highest for those exposed to butanol following negative suggestions, lowest for those exposed to methyl salicylate following positive suggestions" No other interactions

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

						assessed
Devriese et al. (2004) a,e	Non RCT (B+W)	Healthy students (53, U/K, U/K)	Odours	a. Given information about possible health damaging effects of chemical pollution (U/K) b. No information (U/K)	i. Total symptom score (ns) ii. Arousal symptom score (ns) iii. Respiratory symptom score (ns) iv. Cardiac symptom score (ns) v. Tingling symptom score (ns) vi. Unclassified symptom score (ns) vii. Dummy symptom score (ns)	Non assessed
Devriese et al. (2006)	RCT (B+W)	Psychology students (40, U/K, 0.0)	Odour	a. Given information about possible health damaging effects of chemical pollution (20) b. No information (20)	i. Total symptom score (ns) ii. Arousal symptom score (ns) iii. Respiratory symptom score (ns) iv. Cardiac symptom score (ns) v. Tingling symptom score (ns) vi. Unclassified symptom score (ns) vii. Dummy symptom score (ns)	Non assessed
Gavrylyuk et al. (2010) e	RCT (B)	Healthy volunteers (30, 24.9, 32.0)	Saline eye drops	a. Informed of pupil dilation effects (10) b. Informed of pupil constriction effects (10) c. Informed of saline eye drops (10)	i. Frequency of symptoms (a+b>c, "Subjective symptoms, were more frequently reported by participants in the experimental placebo groups than by controls"; a vs b, ns)	N/A
Gibbons et al. (1979) a,e	RCT (B)	Female students (38, U/K, 0.0)	Sham drug	a. Told they were taking Cavanol which would produce some noticeable side effects (19) b. Told they were taking baking soda (19)	i. Perceived arousal (ns) ii. Salient Symptom score (a>b, p<.001) iii. Non salient symptom score (a>b, p<.01)	Symptom suggestion x Self-awareness i. ns ii. Condition a reported less symptoms than condition b when mirror was facing them, (p<.03) iii. ns
Heatherton et al. (1989) e	RCT (B)	Female students (59, U/K, 0.0)	Sham vitamin pill	a. Told vitamin has been reported to make people feel hungry (19) b. Told vitamin has been reported to make people feel full (20) c. Told no further information (20)	i. Hunger ratings (ns)	No significant interactions with participant restraint

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Jaen and Dalton (2014) a,b,e	Non RCT (B)	Asthmatics (17, 38.5, 52.9)	Sham active odour	a. Labelled the odour as therapeutic (9) b. Labelled the odour as asthmogenic (8)	i. Asthma symptom checklist (ns)	N/A
Neukirch and Colagiuri (2014) a,e	RCT (B)	Students with sleep difficulty (91, 21.3, 33.0)	Sham sleep medication	a. Warned about an increase/decrease in appetite and received placebo treatment (24) b. Warned about the side effect but received no treatment (23) c. Not warned about the side effects and received placebo treatment (22) d. Not warned about the side effects and received no treatment (22)	i. Free reporting of side effects (ns) ii. Change in appetite (ns) iii. Change in appetite consistent with warning (a, p=.008) iv. Severity of changes in appetite (lower for those warned about a decrease compared to those warned of an increase, p=.02, or those not warned, p=.012) v. c vs d (change in appetite, ns; direction of change, ns; severity of change, ns)	N/A
Pennebaker and Skelton (1981) e	RCT (B)	Students (38, U/K, 31.6)	Ultrasonic noise	a. Told it would increase skin temperature (13) b. Told it would decrease skin temperature (12) c. Told it would have no effect on skin temperature (13)	i. Perceptions of skin temperature (a>b, p<.001; a>c, p<.05; b vs c, ns)	N/A
Put et al. (2004) a,b,c,e	W	Asthma patients (32, 40.0, 50.0)	Sham inhaler	a. Told it would have no effect on breathing b. Told it was a bronchoconstrictor c. Told it was a bronchodilator	i. Obstruction (b>c, p<.01; a vs b, a vs c, ns) ii. Dyspnea (b>c, p<.01; a vs b, a vs c, ns) iii. Fatigue (ns) iv. Hyperventilation (ns) v. Anxiety (ns) vi. Irritation (ns)	Symptom suggestion x Negative affect i+ii. Higher after condition b than c for participants with high negative affect (i, p<.01; ii, p<.05) iii-vi. ns No significant interactions with social desirability
Read and Bohr (2014) a,b,c,e	Non RCT (B)	Volunteers without photosensitive epilepsy (177, 25.3, U/K)	Sham 3D TV	a. Told it was 3D and wore passive 3D glasses (22) b. Told it was 3D and wore active no shuttering 3D glasses (33) c. Told it was 2D and did not wear glasses (122)	i. Symptom checklist score (a+b>c, p=.03; a vs b, ns)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Schweiger and Parducci (1981) e	RCT B)	Students (34, U/K, 52.9)	Sham electric current	a. Told a low current would be delivered, too mild to be felt but had produced mild headaches in the past (17) b. Told current would be too weak to be felt, but some people develop mild headaches as a side effect (17)	i. Headache pain rating (ns)	N/A
Walach et al. (2002)	RCT (B)	Coffee drinkers (159, 25.5, 58.0)	Sham coffee	a. Received an information leaflet describing the pharmacological effects of caffeine (U/K) b. Received no further information (U/K)	i. General wellbeing score (ns)	No significant interaction with likelihood suggestion
Winters et al. (2003) e	Non RCT (B+W)	18-30 year olds (32, U/K, 15.6)	Odour	a. Given leaflet describing widespread chemical pollution of the environment is a potential cause of multiple chemical sensitivity (16) b. No information given (16)	i. Total symptom score (ns) ii. Arousal symptom score (ns) iii. Respiratory symptom score (ns) iv. Cardiac symptom score (ns) v. Tingling symptom score (ns) vi. Unclassified symptom score (ns) vii. Dummy symptom score (ns)	Symptom suggestions x Conditioning i+iii+iv+vi. Higher following odour which was paired with CO2 when given symptom suggestions (i, p<.01; iii+iv+vi, p<.05) ii+v+vii. ns No other interactions assessed
Wise et al. (2009) c	RCT (B)	Patients with poor asthma control (241, 39.0, 29.5)	Sham asthma drug	a. Emphasized benefit of treatment and described potential side effects (121) b. Expressed uncertainty about improvement following treatment and did not describe potential side effects (120)	i. Headaches (a>b, p=.03) ii. Lethargy (ns) iii. Gastrointestinal distress (ns) iv. Fever (ns) v. Rhinitis (ns) vi. Cough (ns) vii. Flu (ns) viii. Skin rash (ns)	N/A
Witthöft and Rubin (2013)	RCT (B)	Adult English speakers (147, 29.8, 32.7)	Sham EMF	a. Watched a documentary concerning the potential adverse health effects of Wi-Fi (76) b. Watched a BBC News report concerning the security of the internet	i. Total symptom score (ns) ii. Anxiety (ns) iii. Head and concentration (ns) iv. Tingling sensations (ns)	Symptom suggestion x Anxiety i+iii. Increased in people with high anxiety who were in

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

				and mobile phone data (71)		condition a (i, p=.008; iii, p<.001) ii+iv. ns No significant interactions with age, gender, level of education or personality
Zimmerman n-Viehoff et al. (2013) b,e	RCT (B)	Health Caucasians (92, 24.5, 41.3)	Sham arousal oral spray	a. Told it contained a drug to increase BP (33) b. Told it contained a drug to decrease BP (29) c. Told it was a placebo (30)	i. Perceived drug effect (a>c, p=.04; b>c, p=.003; a vs b, ns)	N/A

Note: RCT = Randomised Controlled Trial, Non RCT = Non randomised controlled trial, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, italicised = not directly given but has been extrapolated from the available data, **a** = high risk random sequence generation bias, **b** = high risk allocation concealment bias, **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 11. Miscellaneous risk factors for symptom reporting in response to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Manipulation	Experimental conditions (n)	Main effect on symptoms measured	Interaction(s) with other risk factors
Faasse et al. (2013) b,c,e	RCT (B)	Healthy students (60, 19.4, 43.5)	Sham anti-anxiety tablet	Verbal suggestion (Brand)	a. Branded reformulation change (20) b. Generic reformulation change (20) c. No change (20)	i. Number of expected symptoms (b>c, p=.03; a vs b, ns; a vs c, ns) ii. Number of unexpected symptoms (ns)	N/A
Geers, Weiland, et al. (2005) e	RCT (B)	Healthy students (57, U/K, 35.1)	Sham caffeine pill	Cooperation prime	a. Given a scrambled sentence test with a cooperation prime (U/K) b. Given a scrambled sentence test with a neutral prime (U/K)	i. Caffeine symptom questionnaire score (ns)	Cooperation prime x Arousal suggestion i. Higher in condition a than b when told they were given coffee (p=.02) No significant interactions with caffeine consumption
Jensen and Karoly (1991) e	RCT (B+W)	Students (86, U/K, 45.3)	Sham sedative pill	Verbal suggestion (social desirability)	a. Type B personality is more positive than type A. Type B have been shown to respond more to pills (43) b. Relationship between type A and B personality and response to pills is very weak (43)	i. General placebo response rating (a>b, p<.05) ii. Adjective symptom checklist score (a>b, p<.05)	No significant interactions with dose or gender
Walach et al. (2001)	RCT (B)	Coffee drinkers (157, 28.1, 34.0)	Sham coffee	Experimenter expectancy	a. Experimenter told the physiological effects from a caffeine placebo are real (proplacebo) (U/K) b. Experimenter told the effects of caffeine placebos are just due to artefacts (antiplacebo) (U/K)	i. General wellbeing score (ns)	No significant interactions with likelihood suggestion
Van Diest et al. (2006) e	RCT (B+W)	Students (28, U/K, 21.4)	Odours	Type of breathing	a. Test odours given with normocapnic breathing trial (U/K) b. Test odours given with spontaneous breathing (U/K)	i. Paresthesia (ns) ii. Cerebral (ns) iii. Cardiac (ns) iv. Gastrointestinal (ns) v. Respiratory (a>b, no statistics given) vi. Anxiety (ns)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

						vii. Neuropsychological (ns) viii. Unclassified (ns) ix. Dummy (ns) x. Feeling unreal (ns)	
Devriese et al. (2000) a,e	Non RCT (B+W)	Healthy students (56, U/K, 41.1)	Odours	Timing	a. Test phase immediately after conditioning trials (28) b. Test phase one week after conditioning trials (28)	i. Total symptom score (ns) ii. Arousal (a>b, no statistics given) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	Not assessed

Note: RCT = Randomised Controlled Trial, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, italicised = not directly given but has been extrapolated from the available data, **a** = high risk random sequence generation bias, **b** = high risk allocation concealment bias **c** = high risk blinding of participants and personnel bias, **e** = did not mention an a priori sample size calculation, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 12. Demographic predictors of symptom reporting to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Risk factor(s)	Symptoms measured	Interaction(s) with other risk factors
Angelucci and Pena (1997) e	RCT (B)	Student caffeine consumers (148, U/K, 23.0)	Sham Coffee	1. Gender (F,M)	I. Stimulation/alertness (ns) ii. Anxiety/irritability (ns) iii. Subjective symptoms (ns)	Not assessed
Casper et al. (2001) e	P	Nonpsychotic major depressive patients (876, U/K, 42.8)	Sham fluoxetine treatment	1. Gender (F,M)	i. Number of people reporting one or more symptoms (F>M, p<.01) ii. Specific symptoms (F reported a higher incidence of pain in general, chest pain, infections, accidental injuries, nausea, increased appetite, and nervousness, ps<.043. M were more likely to report somnolence, tremor, and asthma, ps<.048)	Not assessed
de la Cruz et al. (2010) e	P	Patients with cancer related fatigue (105, U/K, 40.0)	Sham treatment	1. Age 2. Education level	i. Nausea (1 ns; higher 2 increases i, p=.05)	Not assessed
Drici et al. (1995) b,e	P	Healthy volunteers (52, 23.5, 50.0)	Sham paracetamol eye drop	1. Employment	i. Subjective side effect rating scale (ns)	Not assessed
Geers et al. (2011) e	RCT (B)	Healthy students (102, 20.5, 21.6)	Sham caffeine capsule	1. Age 2. Gender	i. Placebo response index- anxious, sluggish, energized, calm, irritated, lazy, relaxed, and excited (1 ns; 2 ns)	No significant interactions with likelihood suggestion
Geers, Helfer, et al. (2005) e	RCT (B)	Healthy students (54, 21.0, 29.6)	Sham over-the-counter pill	1. Age 2. Gender	i. Affect questionnaire - Anxiety, restlessness, relaxed, irritable, and perspiration (1 ns; 2 ns)	No significant interactions with self-awareness, likelihood suggestion or optimism
Goetz et al. (2008) e	P	Parkinson's patients with dyskinesia (484, U/K, U/K)	Sham medication	1. Age 2. Gender	i. UPDRS score worsening (1 ns; 2 ns)	Not assessed
Harrell and Juliano (2012) e	RCT (B)	Adult smokers (43, 28.7, 67.4)	Sham cigarette	1. Gender	i. Cigarette evaluation scale - cravings (ns)	No significant interaction with performance suggestion

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Jensen and Karoly (1991) e	RCT (B+W)	Students (86, U/K, 45.3)	Sham sedative pill	1. Gender	i. General placebo response rating (ns) ii. Adjective symptom checklist score (ns)	No significant interaction with dose or social desirability
Liccardi et al. (2004) b,e	P	Patients with ADRs (600, 42.0, 30.3)	Sham allergen pill	1. Gender	i. Number of people reporting reactions (F>M, p=.01)	Not assessed
Lombardi et al. (2008) a,d,e	P	Patients with ADRs (435, 39.7, 32.0)	Sham allergen pill	1. Age 2. Gender	i. Number of recorded symptoms (1 ns; 2 ns)	Not assessed
Lorber et al. (2007) e	RCT (B)	Students without upper respiratory conditions (86, U/K, 40.7)	Sham environmental toxin	1. Gender	i. Verbal symptom ratings: specified-headache, nausea, itchy skin, drowsiness and additional - watery eyes, scratchy throat, chest tightness, and breathing difficulty (ns)	Gender x Observation i. Higher in observation compared to no observation conditions when participants are female (p<.05)
Mazzoni et al. (2010) e	RCT (B)	Healthy students (120, 20.7, 50.0)	Sham environmental toxin	1. Gender	i. Verbal symptom ratings: specified-headache, nausea, itchy skin, drowsiness and additional - watery eyes, scratchy throat, chest tightness, and breathing difficulty (ns)	Not assessed
Papoiu et al. (2011) e	RCT (W)	Healthy volunteers and patients with atopic dermatitis (25, U/K, 44.0)	Sham histamine	1. Gender	i. Average itch intensity rating (ns) ii. Scratching behaviour (ns)	No significant interactions with social observation
Read and Bohr (2014) e	Non RCT (B)	Volunteers without photosensitive epilepsy (177, 25.3, U/K)	Sham 3D TV	1. Gender	i. Symptom checklist score (ns)	Not assessed
Strohle (2000) e	P	Healthy adults and patients with panic disorder (U/K, 33.5, 56.6)	Sham panic disorder trigger	1. Gender	i. Acute panic inventory rating scale (Healthy adults, ns; patients F>M, p<.05)	Gender x Condition i. Increases for females with panic disorder (p<.05)
Szemerszky et al. (2010) e	W	Healthy students (40, 22.8, 27.5)	Sham EMF	1. Gender	i. Overall symptom score (weak suggestion, F>M, p<.05; strong suggestion, ns)	Not assessed
Van den Bergh et al. (1997) a,e	Non RCT (B+W)	Psychosomatic patients (28, 36.0, 50.0)	Odours	1. Gender	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

					v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	
Van den Bergh et al. (1998) e	RCT (B+W)	Healthy adults (56, 42.5, 50.0)	Odours	1. Gender	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	Not assessed
Witthöft and Rubin (2013)	RCT (B)	Adult English speakers (147, 29.8, 32.7)	Sham EMF	1. Age 2. Gender 3. Level of education	i. Total symptom score (1 ns; 2 ns; 3 ns) ii. Anxiety (1 ns; 2 ns; 3 ns) iii. Head and concentration (1 ns; 2 ns; 3 ns) iv. Tingling sensations (1 ns; 2 ns; 3 ns)	No significant interactions with personality, anxiety or symptom suggestion

Note: RCT = Randomised Controlled Trial, Non RCT = Non randomised controlled trial, P = Prospective design, B = Between subjects design, W = Within subjects design, U/K= Unknown, italicised = not directly given but has been extrapolated from the available data, F = Female, M = Male, ns = non-significant, Not assessed = did not assess interactions with another risk factor, **a** = high risk for selection bias, **b** = high risk for confounding factors, **d** = high risk for low generalizability, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 13. Clinical characteristics as predictors of symptom reporting to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Risk factor(s)	Symptoms measured	Interaction(s) with other risk factors
Andre-Obadia et al. (2011) e	RCT (W)	Chronic neuropathic pain patients (45, 55.0, 37.8)	Sham rTMS	1. Pain ratings	i. Pain rating (ns)	No significant interaction with prior experience
Bogaerts et al. (2010) e	P	Patients with medically unexplained dyspnea and healthy controls (58, U/K, 0.0)	Breathing trial with room air	1. Clinical condition	i. Dyspnea score (patients reported higher scores than controls, $p<.05$)	Not assessed
Casper et al. (2001) e	P	Nonpsychotic major depressive patients (876, U/K, 42.8)	Sham fluoxetine treatment	1. Depression severity	i. Number of people reporting one or more symptoms (ns)	Not assessed
Danker-Hopfe et al. (2010)	P	German villages with weak RF-EMF sources (397, U/K, 49.1)	Sham EMF	1. Sleep quality	i. Subjective sleep quality (lower 1 decreased i, $p<.001$)	Not assessed
de la Cruz et al. (2010) e	P	Patients with cancer related fatigue (105, U/K, 40.0)	Sham treatment	1. Cancer performance status 2. Well being 3. Cognitive status 4. Nausea 5. Sleep quality 6. Anxiety symptoms	i. Dizziness (worse 1 increased i, $p=.03$) ii. Insomnia (lower 2 increased ii, $p=.01$; higher 3 increased ii, $p=.01$; higher 4 increased ii, $p=.04$; lower 5 increased ii, $p=.04$) iii. Nausea (higher 4 increased iii, $p=.004$) iv. Restlessness (higher 6 increased iv, $p=.002$)	Not assessed
De Peuter et al. (2005) e	RCT (W)	Asthma patients and healthy controls (40, 23.9, 52.5)	Sham inhaler	1. Clinical condition	i. Total symptom score (ns) ii. Obstruction (ns) iii. Dyspnea (ns) iv. Fatigue (ns) v. Hyperventilation (Asthma patients scored higher than healthy controls, $p<.05$) vi. Anxiety (ns) vii. Irritation (ns)	No significant interaction with association. No other interactions assessed.
Flaten et al. (2003) e	W	Coffee drinkers (20, U/K, 50.0)	Sham coffee	1. Symptoms	i. Alertness (ns) ii. Contentedness (ns) iii. Calmness (for 1 cup $r=-.69$, $p<.01$; for 2 cups $r=-.71$, $p<.01$) iv. Arousal (for 1 cup $r=-.76$, $p<.01$; for 2 cups $r=-.6$, $p<.01$)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

					v. Stress (for 1 cup $r=-.67$, $p<.005$; for 2 cups, ns)	
Goetz et al. (2008) e	P	Parkinson's patients with dyskinesia (484, U/K, U/K)	Sham medication	1. Dyskinesia severity 2. UPDRS motor score 3. Daily L-dopa dose 4. Dyskinesia duration 5. Adverse events 6. Adverse event severity	i. UPDRS score worsening (lower 1 increased i, $p<.0001$; 2 ns; 3 ns; 4 ns; 5 ns and 6 ns)	Not assessed
Lombardi et al. (2008) a,d,e	P	Patients with ADRs (435, 39.7, 32.0)	Sham allergen pill	1. Atopic status 2. Previous reaction severity 3. Type of previous reaction	i. Recorded symptoms (1 ns; 2 ns; 3 ns)	Not assessed
Mrna and Skrivaneck (1985) e	W	Healthy volunteers (21, 17.0, 47.6)	Sham arousal drug	1. Response to other placebo	i. Drug effect questionnaire score ($K=.67$)	Not assessed
Nevelsteen et al. (2007) e	RCT (B)	Healthy males (59, 48.4, 100.0)	Sham magnetic field	1. Depression	i. Subjective vigilance feelings (ns) ii. Profile of mood states (ns) iii. 24 Physical symptoms scale (ns)	Not assessed
Papoiu et al. (2011) e	RCT (W)	Healthy volunteers and patients with atopic dermatitis (25, U/K, 44.0)	Sham histamine	1. Clinical condition	i. Average itch intensity rating (higher in itch video compared to neutral video for patients, $p=.027$; healthy volunteers, ns) ii. Itching behaviour (patients scratched more frequently in areas beyond the itch site, $p=.001$, compared to healthy volunteers when watching the itch video)	Not assessed
Strohle (2000) e	P	Healthy adults and patients with panic disorder (U/K, 33.5, 56.6)	Sham panic disorder trigger	1. Clinical condition	i. Acute panic inventory rating scale (patients scored higher than healthy volunteers, $p<.05$)	Condition x Gender i. Increased in Females with panic disorder ($p<.05$)
Szemerszky et al. (2010) e	W	Healthy students (40, 22.8, 27.5)	Sham EMF	1. IEI EMF score	i. Overall symptom score (for weak suggestion $r=.46$, $p<.01$; for strong suggestion; $r=.48$, $p<.01$; regression, ns)	Not assessed

Note: RCT = Randomised Controlled Trial, P = Prospective design, B = Between subjects design, W = Within subjects design, U/K= Unknown, italicised = not directly given but has been extrapolated from the available data, UPDRS = Unified Parkinson's disease rating scale, IEI-EMF = idiopathic environmental intolerance attributed to electromagnetic fields, ns = non-significant, Not assessed = did not assess interactions with another risk factor, **a** = high risk for selection bias, **d** = high risk for low generalizability, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 14. Expectations as predictors of symptom reporting to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Risk factor(s)	Symptoms measured	Interaction(s) with other risk factors
Angelucci and Pena (1997) e	RCT (B)	Student caffeine consumers (148, U/K, 23.0)	Sham coffee	1. Expectations of the effect of coffee	I. Stimulation/alertness (ns) ii. Anxiety/irritability (ns) iii. Subjective symptoms (ns)	Not assessed
Bayer et al. (1998) a,e	RCT (B)	Job seekers (62, U/K, 82.0)	Sham electrical shock	1. Believed what they were told	i. Subjects reporting pain (ns)	Not assessed
De Peuter et al. (2005) e	RCT (W)	Asthma patients and healthy controls (40, 23.9, 52.5)	Sham inhaler	1. Symptom expectations	i. Total symptom score (overall, $r=0.52$, $p<.001$; control, ns; asthma patients, $r=0.69$, $p<.001$)	Not assessed
Fillmore and Vogel-Sprott (1992) e	P	Male students (56, U/K, 100.0)	Sham coffee	1. Symptom expectations	i. Alertness score (higher 1 increased i, $p<.001$) ii. Tension score (higher 1 increased ii, $p<.001$)	N/A
Flaten et al. (2003) e	W	Coffee drinkers (20, U/K, 50.0)	Sham coffee	1. Symptom expectations	i. Alertness (for 1 cup $r=.63$, $p<.005$; for 2 cups $r=.76$, $p<.001$) ii. Discontentedness (for 1 cup $r=.57$, $p<.01$; for 2 cups, ns) iii. Calmness (ns) iv. Arousal (ns) v. Stress (ns)	Not assessed
Köteles and Babulka (2014) a,d,e	P	Adult volunteers (33, 37.7, 15.2)	3 types of Essential oils (Randomised to 1)	1. Symptom expectations	i. Perceived change in alertness: rosemary oil (ns), lavender oil (higher 1 increased i, $p<.001$), eucalyptus oil (ns) ii. Perceived change in heart rate: rosemary oil (ns), lavender oil (ns), eucalyptus oil (ns) iii. Perceived change in BP: rosemary oil (ns), lavender oil (ns), eucalyptus oil (ns)	Not assessed
Link et al. (2006) a,b,c,d,e	P	Students (36, 22.7, 44.0)	Sham herbal supplement	1. Belief they had taken active supplement	i. Number of symptoms reported (those who thought they had taken the active supplement reported more symptoms than those who thought they had taken the placebo, $p=.003$)	Not assessed
Molcan and et al. (1982)	P	Medical students (48, U/K, 52.1)	Sham arousal pill	1. Symptom expectations	i. Symptom scale score (ns)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

b,e						
Szemerszky et al. (2010)	W	Healthy students (40, 22.8, 27.5)	Sham EMF	1. Symptom expectations	i. Overall symptom score (for weak EMF $r=.49$, $p<.01$; for strong EMF $r=.42$, $p<.01$; regression ns)	Not assessed
e						
Vase et al. (2013)	P	Patient with pain due to tooth removal (U/K, 25.5, 47.5)	Sham acupuncture	1. Symptom expectations	i. Pain intensity (higher 1 increased i, $p=.001$) ii. Pain unpleasantness (higher 1 increased ii, $p<.001$)	Not assessed
e						
Walach et al. (2001)	RCT (B)	Coffee drinkers (157, 28.1, 34.0)	Sham coffee	1. General expectations about coffee on wellbeing 2. Subjective probability of receiving coffee	i. General wellbeing score (1 ns; 2 ns)	Not assessed
Walach and Schneider (2009)- exp 1	RCT (B)	Healthy adults coffee drinkers (60, 32.3, 23.3)	Sham caffeine beverage	1. Symptom expectations	i. Mean change in mood (ns) ii. Mean change in calmness (ns) iii. Mean change in alertness (ns)	Not assessed
Walach and Schneider (2009)- exp 2	RCT (B)	Healthy adults coffee drinkers (30, 29.9, 33.3)	Sham caffeine beverage	1. Symptom expectations	i. Mean change in mood (ns) ii. Mean change in calmness (ns) iii. Mean change in alertness (ns)	Not assessed

Note: RCT = Randomised Controlled Trial, P = Prospective design, B = Between subjects design, W = Within subjects design, U/K= Unknown, italicised = not directly given but has been extrapolated from the available data, ns = non-significant, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed, **a** = high risk for selection bias, **b** = high risk for confounding factors, **c** = high risk for insufficient follow-up, **d** = high risk for low generalizability, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 15. Anxiety as a predictor of symptom reporting to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Risk factor(s)	Symptoms measured	Interaction(s) with other risk factors
Angelucci and Pena (1997) e	RCT (B)	Student caffeine consumers (148, U/K, 23.0)	Sham coffee	1. State and trait anxiety	I. Stimulation/alertness (ns) ii. Anxiety/irritability (higher 1 increased ii, p<.0001) iii. Subjective symptoms (ns)	Not assessed
Bogaerts et al. (2010) e	P	Patients with medically unexplained dyspnea and healthy controls (58, U/K, 0.0)	Breathing trial with room air	1. State anxiety	i. Dyspnea score (ns)	Not assessed
Danker-Hopfe et al. (2010)	P	Villages in Germany with weak RF-EMF sources (397, U/K, 49.1)	Sham EMF	1. General fear/anxiety towards risks of RF-EMF 2. Fear/anxiety towards base station	i. Subjective sleep quality (1 ns; higher 2 decreased i, p<.05)	Not assessed
Link et al. (2006) a,b,c,d,e	P	Students (36, 22.7, 44.0)	Sham herbal supplement	1. State anxiety	i. Number of symptoms reported (ns)	Not assessed
Molcan and et al. (1982) b,e	P	Medical students (48, U/K, 52.1)	Sham arousal pill	1. State anxiety 2. Trait anxiety	i. Symptom scale score (1 ns; 2 ns)	Not assessed
Nevelsteen et al. (2007) e	RCT (B)	Healthy males (59, 48.4, 100.0)	Sham magnetic field	1. State anxiety 2. Trait anxiety	i. Subjective vigilance feelings (1 ns; 2 ns) ii. Profile of mood states (1 ns; 2 ns) iii. 24 Physical symptoms scale (higher 1 increased iii, p<.001; 2 ns)	Not assessed
Szemerszky et al. (2010) e	W	Healthy students (40, 22.8, 27.5)	Sham EMF	1. State anxiety	i. Overall symptom score (ns)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Van den Bergh et al. (1997) a,e	Non RCT (B+W)	Psychosomatic patients (28, 36.0, 50.0)	Odours	1. State and trait anxiety	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	Not assessed
Witthöft and Rubin (2013)	RCT (B)	Adult English speakers (147, 29.8, 32.7)	Sham EMF	1. State anxiety	i. Total symptom score (ns) ii. Anxiety (ns) iii. Head and concentration (ns) iv. Tingling sensations (ns)	Anxiety x Symptom suggestion i+iii. Increased in people with high levels of anxiety who were in Wi-Fi group (i, p=.008; iii, p<.001). ii+iv. ns No significant interactions with age, gender, level of education or personality

Note: RCT = Randomised Controlled Trial, P = Prospective design, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, Not assessed = did not assess interactions with another risk factor, **a** = high risk for selection bias, **b** = high risk for confounding factors, **c** = high risk for insufficient follow-up, **d** = high risk for low generalizability, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 16. Personality as a predictor of symptom reporting to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Risk factor(s)	Symptoms measured	Interaction(s) with other risk factors
Angelucci and Pena (1997) e	RCT (B)	Student caffeine consumers (148, U/K, 23.0)	Sham coffee	1. Suggestibility	I. Stimulation/alertness (ns) ii. Anxiety/irritability (ns) iii. Subjective symptoms (ns)	Not assessed
Bogaerts et al. (2010) e	P	Patients with medically unexplained dyspnea and healthy controls (58, U/K, 0.0)	Breathing trial with room air	1. Negative affect	i. Dyspnea score (ns)	Not assessed
Davis et al. (1995) a,d,e	P	Healthy Adults (27, U/K, 55.6)	Sham anti-depressant pill	1. Neuroticism 2. Somatosensory amplification	i. Side effect checklist (1, “significant positive correlation”; 2, ns)	Not assessed
De Peuter et al. (2005) e	RCT (W)	Asthma patients and healthy controls (40, 23.9, 52.5)	Sham inhaler	1. Negative affect	i. Total symptom score (ns) ii. Obstruction (ns) iii. Dyspnea (ns) iv. Fatigue (ns) v. Hyperventilation (ns) vi. Anxiety (ns) vii. Irritability (higher 1 increased vii, p<.05)	Not assessed
De Peuter et al. (2007) e	P	Asthma patients (30, 38.0, 26.7)	Sham histamine inhalation	1. Negative affect	i. Obstruction (higher 1 increased i, p<.05) ii. Dyspnea (ns) iii. Fatigue (higher 1 increased iii, p<.001) iv. Hyperventilation (ns) v. Anxiety (ns) vi. Irritability (higher 1 increased vi, p<.001)	None
Devriese et al. (2000) a,e	Non RCT (B+W)	Healthy students (56, U/K, 41.1)	Odours	1. Negative affect	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns)	Conditioning x Odour x Negative affect i+iii. Higher in response to odour paired with CO2 but only when the odour was ammonia and

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

					vi. Unclassified (ns) vii. Dummy (ns)	participants had high negative affect (p<.05) ii, iv-vii. ns Odour x Negative affect x Generalisation i+iii. Higher in response to butyric and acetic acid than citric aroma when ammonia was paired with CO2 and participants had high negative affect (p<.05) iv. Higher in response to butyric acid than acetic acid or citric aroma when ammonia was paired with CO2 and participants had high negative affect (p<.05) ii+ v-vii. Ns No other interactions assessed
Devriese et al. (2004) a,e	Non RCT (B+W)	Healthy students (53, U/K, U/K)	Odours	1. Negative affect	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	Non assessed
Drici et al. (1995) b,e	P	Healthy volunteers (52, 23.5, 50.0)	Sham paracetamol eye drop	1. Type A Personality 2. Type B Personality	i. Subjective side effect reports (1>2, p=.03)	Not assessed
Geers, Helfer, et al. (2005) e	RCT (B)	Healthy students (54, 21.0, 29.6)	Sham over-the-counter pill	1. Optimism-pessimism	i. Affect questionnaire - Anxiety, restlessness, relaxed, irritable, and perspiration (ns)	Optimism x Likelihood suggestion i. Increased score for pessimists than optimists in deceptive group than the control (p<.05), no significant difference between conditional group and control. No significant interaction with age, gender or self-awareness
Heatherton et al. (1989)	RCT (B)	Female students (59, U/K, 0.0)	Sham vitamin pill	1. Restraint	i. Hunger ratings (ns)	No significant interaction with symptom suggestion

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

e						
Link et al. (2006) a,b,c,d,e	P	Students (36, 22.7, 44.0)	Sham herbal supplement	1. Social desirability	i. Number of symptoms reported (ns)	Not assessed
Mazzoni et al. (2010) e	RCT (B)	Healthy students (120, 20.7, 50.0)	Sham environmental toxin	1. Openness 2. Conscientiousness 3. Extraversion 4. Agreeableness 5. Neuroticism	i. Verbal symptom reports (1 ns; 2 ns; 3 ns; 4 ns; 5 ns)	Not assessed
Nevelsteen et al. (2007) e	RCT (B)	Healthy males (59, 48.4, 100.0)	Sham magnetic field	1. Positive affect 2. Negative affect 3. Sensitivity to anxiety 4. Vigilance	i. Subjective vigilance feelings (1 ns; 2 ns; 3 ns; 4 ns) ii. Profile of mood states (1 ns; 2 ns; 3 ns; 4 ns) iii. 24 Physical symptoms score (higher 1 decreased ii, p<.001; 2+3 ns; higher 4 increased iii, p<.001)	Not assessed
Put et al. (2004) a,b,c,e	W	Asthma patients (32, 40.0, 50.0)	Sham inhaler	1. Negative affect 2. Social desirability	i. Obstruction (higher 1 increased i, p<.01; 2 ns) ii. Dyspnea (higher 1 increased ii, p<.01; 2 ns) iii. Fatigue (higher 1 increased iii, p<.05; 2 ns) iv. Hyperventilation (higher 1 increased iv, p<.05; 2 ns) v. Anxiety (higher 1 increased v, p<.05; 2 ns) vi. Irritation (higher 1 increased vi, p<.05; 2 ns)	Symptom suggestion x negative affect i+ii. Higher after bronchoconstriction than bronchodilator suggestions for those with high negative affect (i, p<.01; ii, p<.05) iii-vi. ns No significant interactions with social desirability
Slanska et al. (1974) a,e	Non RCT (B)	Medical students (33, U/K, U/K)	Indifferent salt solution	1. Stability – instability 2. Activity – passivity 3. Submissive – dominance 4. Rationality – sensuousness 5. Introversion - extraversion	i. Perceived effect ("correlated with 1+3 personality characteristics. Whereby significant changes in the direction of suggestion mainly affecting individuals unstable and submissive")	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Stegen et al. (1998) e	RCT (W)	Healthy psychology students (72, U/K, 48.6)	Breathing trial with room air	1. Negative affect	i. Total symptom score (ns) ii. General arousal (ns) iii. Respiration (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Gastrointestinal (ns) viii. Dizziness (ns)	Prior experience x Negative affect i+ii. Higher in participants scoring high on negative affect compared to low when room air trial was before CO2 trial (i, p<.001; ii, p<.005) iii-viii. ns
Stegen et al. (2000) a,b,d,e	P	Healthy psychology students (44, U/K, 27.3)	Breathing trial with room air	1. Negative affect 2. Social desirability	i. Somatic experience intensity (higher 1 increased i, p<.01; 2 ns) ii. Unpleasantness (1 ns; 2, ns) iii. General arousal (1 ns; 2 ns) iv. Respiration (higher 1 increased iv, p<.05; 2 ns) v. Cardiac (1 ns; 2 ns) vi. Tingling (higher 1 increased vi, p<.05; 2 ns) vii. Gastrointestinal (1 ns; 2 ns) viii. Unclassified sensations (higher 1 increased viii, p<.05; 2 ns) ix. Dummy sensations (1 ns; 2 ns)	Not assessed
Sullivan et al. (2008) e	P	Patients with neuropathic pain (24, 54.7, 62.5)	Sham cream treatment	1. Pain catastrophising	i. Side effects reported (r=0.29, p<.05)	N/A
Szemerszky et al. (2010) e	W	Healthy students (40, 22.8, 27.5)	Sham EMF	1. Dispositional optimism 2. Somatisation 3. Somatosensory amplification 4. Motivation	i. Overall symptom score (1+4 significant negative correlation; 2+3 significant positive correlation for weak and strong; regression of 2,3,4 was significant for weak suggestion, only 2+4 were significant for strong suggestion)	Not assessed
Van den Bergh et al. (1995) a,e	Non RCT (B+W)	Healthy students (28, U/K, 50.0)	Odours	1. Negative affect	i. Total symptom score (ns) ii. Arousal symptom score (ns) iii. Respiratory symptom score (ns) iv. Cardiac symptom score (ns) v. Tingling symptom score (ns) vi. Unclassified symptom score (ns)	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Van den Bergh et al. (1997) a,e	Non RCT (B+W)	Psychosomatic patients (28, 36.0, 50.0)	Odours	1. Blunting behaviour	i. Total symptom score (Higher 1 increased i, p<.002) ii. Arousal symptom score (ns) iii. Respiratory symptom score (ns) iv. Cardiac symptom score (ns) v. Tingling symptom score (ns) vi. Unclassified symptom score (Higher 1 increased vi, p<.005) vii. Dummy symptom score (ns)	Not assessed
Witthöft and Rubin (2013)	RCT (B)	Adult English speakers (147, 29.8, 32.7)	Sham EMF	1. Perceived sensitivity to EMF 2. Modern health worries 3. Somatisation 4. Somatosensory amplification	i. Total symptom score (1 ns; higher 2 increased i, p<.001; 3 ns; higher 4 increased i, p=.046) ii. Anxiety (ns) iii. Head and concentration (1ns; higher 2 increased iii, p<.001; 3 ns; 4 ns) iv. Tingling sensations (ns)	No significant interactions with age, gender, level of education, anxiety or symptom suggestion

Note: RCT = Randomised Controlled Trial, Non RCT = Non randomised controlled trial, P = Prospective design, B = Between subjects design, W = Within subjects design, U/K= Unknown, ns = non-significant, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed, **a** = high risk for selection bias, **b** = high risk for confounding factors, **c** = high risk for insufficient follow-up, **d** = high risk for low generalisability, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Supplementary Table 17. Miscellaneous baseline predictors of symptom reporting to an inert exposure

Reference and quality	Study design	Population (N, Mean age, %Male)	Inert exposure	Risk factor(s)	Symptoms measured	Interaction(s) with other risk factors
Dalton (1999) e	RCT (B)	Healthy volunteers (180, 31.7, 49.4)	Odours	1. Olfactory sensitivity 2. Odour reactivity	i. Symptom reports (1 ns; higher 2 increases i, $R^2=0.74$)	Not assessed
Danker-Hopfe et al. (2010)	P	Villages in Germany with weak RF-EMF sources (397, U/K, 49.1)	Sham EMF	1. Visibility of the base station 2. Preoccupation with EMF	i. Subjective sleep quality (1 ns; 2 ns)	Not assessed
Devriese et al. (2004) a,e	Non RCT (B+W)	Healthy students (53, U/K, U/K)	Odours	1. Perceived cue odour	i. Total symptom score (ns) ii. Arousal symptom score (ns) iii. Respiratory symptom score (ns) iv. Cardiac symptom score (ns) v. Tingling symptom score (ns) vi. Unclassified symptom score (ns) vii. Dummy symptom score (ns)	Perceived cue odour x Odour i. Higher to butyric acid than ammonia when butyric acid was thought to have been paired with CO2 ($p<.05$) ii-vii. ns No other interactions assessed
Geers et al. (2011) e	RCT (B)	Healthy students (102, 20.5, 21.6)	Sham caffeine capsule	1. Average caffeinated beverage consumption 2. Caffeinated beverages consumed so far that day	i. Placebo response index- anxious, sluggish, energized, calm, irritated, lazy, relaxed, and excited (1 ns; 2 ns)	No significant interactions with likelihood suggestion
Geers, Weiland, et al. (2005) e	RCT (B)	Healthy students (57, U/K, 35.1)	Sham caffeine pill	1. Caffeine consumption	i. Caffeine symptom questionnaire score (ns)	No significant interaction with arousal suggestion
Goetz et al. (2008) e	P	Parkinson's patients with dyskinesia (484, U/K, U/K)	Sham medication	1. Geographical site of enrolment 2. Study (1 or 2)	i. UPDRS score worsening (1 ns; 2 ns)	Not assessed
Goldman et al. (1965) e	Non RCT (B)	Male veterans with Schizophrenia (64, 44.0, 100.0)	Sham arousal treatment	1. High regard for hospital medications	i. Reported symptoms identified through interviews (U/K) ii. Increase in reported drug effects identified through interviews	Not assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

					(C=.30, p<.01) iii. Ward activity (U/K)	
Köteles and Babulka (2014) a,d,e	P	Adult volunteers (33, 37.7, 15.2)	3 types of Essential oils (Randomised to 1)	1. Pleasantness of odour	i. Perceived change in alertness: rosemary oil (ns), lavender oil (ns), eucalyptus oil (increased, p<.05) ii. Perceived change in heart rate: rosemary oil (ns), lavender oil (ns), eucalyptus oil (ns) iii. Perceived change in BP: rosemary oil (ns), lavender oil (ns), eucalyptus oil (increased, p<.05)	Not assessed
Liccardi et al. (2004) b,e	P	Patients with ADRs (600, 42.0, 30.3)	Sham allergen pill	1. Hospital centre	i. Number of people reporting reactions (ns)	Not assessed
Mazzoni et al. (2010) e	RCT (B)	Healthy students (120, 20.7, 50.0)	Sham environmental toxin	1. Gender of model	i. Verbal symptom ratings (more symptoms were reported when participant and confederate were the same gender, d=0.24)	Not assessed
Meulders et al. (2010) a,e	Non RCT (B+W)	Healthy adults (58, 22.0, 48.3)	Odours	1. Ability to predict which odour produced the most symptoms	i. Total symptom score (ns) ii. Arousal (ns) iii. Respiratory (ns) iv. Cardiac (ns) v. Tingling (ns) vi. Unclassified (ns) vii. Dummy (ns)	Conditioning x Ability to predict i. Higher in response to odours which had been paired with CO2 compared to room air when participants were able to predict which odour had caused the most symptoms (p<.05) ii-vii. ns Conditioning x Ability to predict x Odour i-iv+vii. Higher in response to butyric acid which had been paired with CO2 compared to room air when participants were able to predict which odour had caused the most symptoms (i-iv, p<.01; vii, p<.05) v+vi. ns No other interactions assessed

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Nevelsteen et al. (2007) e	RCT (B)	Healthy males (59, 48.4, 100.0)	Sham magnetic field	1. Discomfort under the helmet 2. Stress under the helmet 3. Risk perception	i. Subjective vigilance feelings (1 ns; 2 ns; 3 ns) ii. Profile of mood states (1 ns; 2 ns; 3 ns) iii. 24 Physical symptoms scale (higher 1 increased iii, p<.001; 2, ns; 3, ns)	Not assessed
Wendt et al. (2014) e	P	Healthy males (24, 25.0, 100.0)	Sham immuno- suppressive capsule	1. Genes	i. Number of symptoms reported (More side effects in the Val/Val homozygous carriers compared to the Val158/Met158 groups, p<0.001 and the Met158 homozygous groups, p<.01)	N/A

Note: RCT = Randomised Controlled Trial, P = Prospective design, B = Between subjects design, U/K= Unknown, ns = non-significant, Not assessed = did not assess interactions with another risk factor, N/A = no other risk factors assessed, **a** = high risk for selection bias, **b** = high risk for confounding factors, **d** = high risk for low generalizability, **e** = did not mention an a priori sample size calculation

SYSTEMATIC REVIEW OF NOCEBO EFFECT RISK FACTORS

Appendix 1. Data extraction sheet

Reference	Study 1	Study 2	Study 3	...
Country of origin				
Participants				
Age				
Gender				
Inclusion/exclusion criteria				
Sample size (sample size calculation provided?)				
Number of drop outs				
Methods				
Design				
<u>Materials</u>				
Exposure type				
Experimental condition(s)				
<u>Measures</u>				
Baseline risk factor(s)				
Symptoms				
<u>Statistical analysis</u>				
Baseline risk factor(s)				
Condition				
Results				
Effect of condition on factor(s)				
Effect of baseline risk factor(s) on symptoms				
Ethical Issues				
Ethical approval?				
Details of informed consent				